严士健
刘秀芳 著

测度与概率

CEDU YU
GAILU

北京师范大学出版社

测度与概率

严士健　刘秀芳　著

北京师范大学出版社

(京)新登字160号

责任编辑:潘淑琴
封面设计:任　凯

图书在版编目(CIP)数据

测变与概率/严士健,刘秀芳著.-北京:北京师范大学出版社,1994.11

ISBN 7-303-03790-X

Ⅰ.测… Ⅱ.①严…②刘… Ⅲ.①测度论-基本知识②概率论-基本知识 Ⅳ.①O174.12②O211.1

中国版本图书馆CIP数据核字(94)第15129号

北京师范大学出版社出版发行
(100875　北京新街口外大街19号)
丰润印刷厂印刷　全国新华书店经销
开本:850×1168　印张:12.5　字数:311千
1994年11月第1版　1994年11月第1次印刷
印数:1-3000册
定价:12.00元

责任编辑 [illegible]

封面设计 任 [illegible]

内容提要

[illegible]

前言

这是一本为统计与概率专业而写作的教材．考察概率论与数理统计发展过程，测度论不但已经成为概率论，特别是随机过程论的基础；而且对于数理统计来说，很多基本概念和问题，离开了测度论很难说清楚．这与三十年前的情况大不相同，从发展看，这种依赖测度论的趋势将在未来的世纪中大大加强．因此我们认为在统计与概率专业中应该学习以测度论为基础的概率论基础．为此只学经典的实变函数论是不够的，需要学习测度论．而在讲授测度论时，如果充分以经典的“长度”和概率概念为背景，是可以不必先修实变数函数论的．本书就是在这种认识下编写的，其目的是希望它能够作为在不先修实变数函数论的条件下，学习测度论基础和以测度论为基础的概率论的教材．所以我们在测度论的内容之前，编写了集合、势以及距离空间等题材．

介绍距离空间，更重要的是由于现代概率论以及某些数理统计的题材更适宜于以距离可测空间上的概率测度为框架．介绍这方面的必要知识，可以使测度和概率的一些性质在距离空间的基础上讲授，以使学生具有更广泛的基础，便于以后的学习和应用．当然如果学生在学习本书之前，已经学习过距离空间的知识，自然就可以越过这一段或作简单地复习．这种想法来自一个更广泛的考虑，即教材要适应学生的水平，同时也应该适应学科的现代发展．本书除了将测度与概率尽量建立在距离空间的框架上外，对于有关的题材也向这方面作了一些努力．例如，在书中介绍了 Hausdorff 维数的概念；证明了完备可分距离空间的无穷维乘积空间上的 Колмогоров 相容性定理等．

实际情况告诉我们：测度论虽然是概率论理论的必要基础，但是有时出现学了前者而不能很好地应用于后者的情况．因此我们认为在统计与概率专业的教学中，将二者结合讲授是适宜的．这种看法也许得到概率统计的许多同行的赞同，因为我们和王隽骧教授合著的《概率论基础》[YWL] 曾经得到广泛的传播．本书仍然按照这种观点设计．但是上述著作有其不便之处，例如费了很多篇幅讲解多维分布的情形，作为参考材料可以，作为教材则一些技术细节容易分散学生理解方法的实质的努力．因此本书在分布函数决定测度，**L-S** 积分，特征函数等题材上，讲解一维情形，但注意方法的一般性．希望能收到举一反三的效果．

本书虽然是为统计与概率专业而写，但是如果只采用第一至第五章以及第六、七、八章的前一部分，也可作为其它专业的测度论或实变数函数论课的教材或参考书．将有关概率的一些概念作为一般测度的具体例子，也许对学生理解测度论的作用有帮助，而且还可以扩大学生的知识面．

本书经过我们先后在北京师范大学数学系的统计与概率专业试用四次．栗演兵和张余辉两位先后帮助部分地整理笔记，录入软盘，指出错误．对此我们表示深深的感谢；我们还要感谢那些对讲课提出意见的听课同学，他们的意见使得本书有可能更加适用；感谢潘淑琴，何青和韩丽娟对编辑工作的指导．最后我们要特别提出：没有北京师范大学原自然科学处，出版委员会和出版社的支持和慷慨资助，本书将难以出版．对此我们表示衷心地感谢！

作　者

1994 年 11 月于

北京师范大学数学系

目　录

第一章 集合，映射与势

§1. 集合及其运算

1. 一个给定的 **集合** (或称为 **集**) 是指具有某种性质的事物的全体. 组成集合的每个事物称为该集合的 **元素** (或称为 **元**).

设 A 是一个集合， a 是 A 的元素，则记作 $a \in A$(读作 "a 属于 A") 或 $A \ni a$(读作 "A 包含 a"). 若 a 不是 A 的元素，则记成 $a \notin A$.

给出集合时，一定要有明确规则判定一个事物是否是它的元素. 事实上，这不算是定义，"元素" 和 "集合" 是不定义的名词，还有一个不定义的关系 —"属于". 这些是属于集合论的公理化问题. 有些如拓扑方面的书都设一附录说明，这里就不再仔细推敲.

以后一般用大写字母 $A, B, C, \cdots$ 表示集合，用小写字母 $a, b, c, \cdots$ 表示元素，不够用时可加各种足码.

有时用描述法给出集合，即用 $\{\cdots : \cdots\}$(或 $\{\cdots | \cdots\}$) 表示集合. 括号中 ":" (相应地 "|") 号之前表示元的记法，之后表示该集合的元所具有的性质. 例如：设 $\mathbb{R}$ 为实数集， $\{x : x \in \mathbb{R}, 1 < |x| \le 2\}$ 或 $\{x \in \mathbb{R} : 1 < |x| \le 2\}$ 表示绝对值大于 1 而不超过 2 的全体实数作成的集合. 又例如，设 $\mathbb{C}$ 表示复数集， $\{x \in \mathbb{C} : 1 < |x| \le 2\}$ 表示模大于 1 而不超过 2 的全体复数.

有时，用列举法即用 $\{\cdots\}$ 表示集. 括号中 "$\cdots$" 列出了该集合的所有元，例如： $\{2, 3, 4, 5, 8\}$ 表示由 2, 3, 4, 5, 8 组成的集合， $\{u, v\}$ 表示由 u, v 组成的集合.

若集合 A 的元只有有限多个，则称 A 为 **有限集**，否则称为 **无限集**. 不含任何元素的集合称为 **空集**，例如 $\{x \in \mathbb{R} : |x| < 0\}$，$\{x \in \mathbb{R} : x^2 + 1 = 0\}$ 都是空集.

以下用逻辑符号 "$\exists$" 表示存在，"$\forall$" 表示 "对一切" (或 "对任何") 记自然数集 $\{1, 2, \cdots\}$ 为 $\mathbb{N}$，整数集为 $\mathbb{Z}$，非负整数集为 $\mathbb{Z}_+$.

2. 定义 设 A、B 是两个集合：

1) 集合 A **包含** 集合 B (记作 $B \subset A$ 或 $A \supset B$) 是指对一切 $x \in B$ 必有 $x \in A$.

2) 集合 A, B **相等** (记作 $A = B$) 是指：$A \subset B$ 且 $B \subset A$.

若 $A \subset B$，则称 A 为 B 的 **子集**；若还有 $A \neq B$ (即存在 $y \in B$ 使 $y \notin A$)，则称 A 为 B 的 **真子集**，记作 $A \subsetneq B$. 约定空集 $\varnothing$ 是任何集合的子集.

关系 "$\subset$" 具有 (i) 自反性：对任何集合 $A \subset A$；(ii) 传递性：若 $A \subset B, B \subset C$，则 $A \subset C$；(iii) 反对称性：若 $A \subset B$ 且 $B \subset A$，则 $A = B$.

(i), (ii), (iii) 表明集类对关系 "$\subset$" 作成一个偏序集.

3. 集合的运算

为了以后书写方便，我们采用符号 $A := B$ 表示 A 用 B 定义.

设 A、B 是两个给定的集合，定义

$A \cup B := \{x : x \in A$ 或 $x \in B\}$，称为 A, B 的 **并集**；

$A \cap B := \{x : x \in A$ 且 $x \in B\}$，称为 A、B 的 **交集**；

$A \setminus B := \{x : x \in A$ 且 $x \notin B\}$，称为 A, B 的 **差集**，

当 $B \subset A$ 时，$A \setminus B$ 也称为 B 对 A 的 **余集**；如果只考虑某固定集合 Ω 的子集时，将 $\Omega \setminus B = \{x \in \Omega : x \notin B\}$ 称为 B 的 **余集**，记作 B^c.

$A\Delta B = \{ x : x \in A$ 且 $x \notin B$, 或 $x \in B$ 且 $x \notin A \}$, 称为 A, B 的 **对称差**. 实际上, $A\Delta B = (A \setminus B) \cup (B \setminus A)$.

$$\bigcup_{\alpha \in I} A_\alpha := \{ x : \exists\, \alpha \in I,\ 使\ x \in A_\alpha \};$$

$$\bigcap_{\alpha \in I} A_\alpha := \{ x : \forall\, \alpha \in I,\ 均有\ x \in A_\alpha \}.$$

其中 I 为一指标集.

例1 若 $A_n := (n,\ n+1]$, $B_n := (0,\ n]$, $C_n := (n,\ n+1)$, $n = 0, 1, 2, \cdots$, 则

$$\bigcup_{n=0}^{\infty} A_n = (0, \infty); \quad \bigcup_{n=0}^{\infty} B_n = (0, \infty);$$

$$\bigcup_{n=0}^{\infty} C_n = (0, \infty) \setminus \mathbb{N}; \quad \Omega = \bigcup_{x \in \Omega} \{ x \}.$$

例2 若 $A_n := (0,\ 1 + \frac{1}{n})$, $B_n := (0,\ 1 - \frac{1}{n})$, $n = 1, 2, 3, \cdots$, 则

$$\bigcap_{n=0}^{\infty} A_n = (0,\ 1], \quad \bigcup_{n=0}^{\infty} B_n = (0,\ 1).$$

下列各定理的证明是容易的 (请读者作为习题自证).

4. 定理 集合的交、并运算具有下列性质:

1) $A \cup A = A, \quad A \cap A = A;$ **(幂等性)**

2) $A \cup B = B \cup A, \quad A \cap B = B \cap A;$ **(交换律)**

3) $(A \cup B) \cup C = A \cup (B \cup C),$
$(A \cap B) \cap C = A \cap (B \cap C);$ **(结合律)**

4) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C),$
$(A \cap B) \cup C = (A \cup C) \cap (B \cup C).$ **(分配律)**

5. 定理 集合的差运算具有下列性质:

1) $(A \setminus B) \cap C = (A \cap C) \setminus (B \cap C);$

2) $(A \setminus B) \setminus C = A \setminus (B \cup C);$

3) 若 $A, B \subset \Omega$, 则 $A \setminus B = A \cap B^c$;

4) 设 Ω 是任意一个集合, $\{ A_\alpha : \alpha \in I \}$ 是 Ω 的一族子集, 则有

$$\Omega \setminus \Big(\bigcup_{\alpha \in I} A_\alpha \Big) = \bigcap_{\alpha \in I} (\Omega \setminus A_\alpha),$$
$$\Omega \setminus \Big(\bigcap_{\alpha \in I} A_\alpha \Big) = \bigcup_{\alpha \in I} (\Omega \setminus A_\alpha).$$

4) 称为 **de Morgan 法则**.

6. 定义 设 $\{ A_n : n \in \mathbb{N} \}$ 为一集序列, 称

$$\overline{\lim_{n\to\infty}} A_n := \{ x : x \text{ 属于} A_n, n \in \mathbb{N}, \text{中的无穷多个} \}$$

为此**集序列的上极限**. 称

$$\underline{\lim_{n\to\infty}} A_n := \{ x : x \text{ 不属于} A_n, n \in \mathbb{N}, \text{中的有限多个} \}$$

为此**集序列的下极限**. 若

$$\overline{\lim_{n\to\infty}} A_n = \underline{\lim_{n\to\infty}} A_n,$$

则称 $\{A_n\}_{n=1}^\infty$ 的极限存在, 而上式中的共同集合称为此**集序列的极限**, 记作 $\lim_{n\to\infty} A_n$.

若 $A_n \subset A_{n+1}$, $n \in \mathbb{N}$, 则称 $\{A_n\}_1^\infty$ **单调增**, 记作 $A_n \uparrow$; 若 $A_n \supset A_{n+1}$, $n \in \mathbb{N}$, 则称 $\{A_n\}_1^\infty$ **单调减**, 记作 $A_n \downarrow$. 单调增和单调减集序列统称为**单调集列**.

7. 定理 设 $\{ A_n : n \in \mathbb{N} \}$ 为任一集序列,

1) $\overline{\lim_{n\to\infty}} A_n = \bigcap_{n=1}^\infty \bigcup_{k=n}^\infty A_k$, $\underline{\lim_{n\to\infty}} A_n = \bigcup_{n=1}^\infty \bigcap_{k=n}^\infty A_k$;

2) 若 $\{ A_n \}$ 单调, 则 $\lim\limits_{n\to\infty} A_n$ 存在且

$$\lim_{n\to\infty} A_n = \begin{cases} \bigcup_{n=1}^\infty A_n, & \text{若 } \{A_n\} \text{单调增}; \\ \bigcap_{n=1}^\infty A_n, & \text{若 } \{A_n\} \text{单调减}. \end{cases}$$

只需注意：

$$\overline{\lim_{n\to\infty}} A_n = \{ x : \forall n \in \mathbb{N}, \exists k \geq n, \text{使 } x \in A_k \};$$

$$\underline{\lim_{n\to\infty}} A_n = \{ x : \exists n \in \mathbb{N}, \text{使} \forall k \geq n, x \in A_k \}.$$

详细证明留作习题.

8. 例. 设

$$A_n = \begin{cases} B, & \text{当}n\text{是偶数}, \\ C, & \text{当}n\text{是奇数}. \end{cases}$$

则由定理 7 知

$$\overline{\lim_{n\to\infty}} A_n = B \cup C;$$
$$\underline{\lim_{n\to\infty}} A_n = B \cap C.$$

9. 定义 设 Ω 是一给定的非空集，$A \subset \Omega$. 则称函数

$$I_A(x) = \begin{cases} 1, & x \in A; \\ 0, & x \notin A \end{cases}$$

为 A 的 **示性函数**. 有时也记成 χ_A.

对于 Ω 的一切子集来说，I_A 完全刻划 A. I_A 与 A 之间一一对应，它是联系集合与函数的一个重要工具，需要熟练运用. 它与集合的关系为下述定理所述.

10. 定理 给定不空集合 Ω. 设下述集合都是 Ω 的子集，则

1) $A = \Omega \Longleftrightarrow I_A(x) \equiv 1$. $A = \varnothing \Longleftrightarrow I_A(x) \equiv 0$;

2) $A \subset B \Longleftrightarrow I_A(x) \leq I_B(x)$, 对一切 $x \in \Omega$;

3) $I_{\bigcup_{\alpha \in J} A_\alpha}(x) = \max_{\alpha \in J} I_{A_\alpha}(x)$,

$I_{\bigcap_{\alpha \in J} A_\alpha}(x) = \min_{\alpha \in J} I_{A_\alpha}(x)$, 对一切 $x \in \Omega$;

4) $I_{\overline{\lim}_{n\to\infty} A_n}(x) = \overline{\lim_{n\to\infty}} I_{A_n}(x)$, $I_{\underline{\lim}_{n\to\infty} A_n}(x) = \underline{\lim_{n\to\infty}} I_{A_n}(x)$.

因而 $\lim\limits_{n\to\infty} A_n$ 存在 $\Longleftrightarrow \lim\limits_{n\to\infty} I_{A_n}$ 存在 (即对任何 $x\in\Omega$, $\lim\limits_{n\to\infty} I_{A_n}(x)$ 存在), 且 $I_{\lim\limits_{n\to\infty} A_n} = \lim\limits_{n\to\infty} I_{A_n}$.

证明 1)、2)、3) 留作习题，我们只证 4):

$$I_{\overline{\lim\limits_{n\to\infty}} A_n}(x) = I_{\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k}(x) = \min_{n\geq 1} I_{\bigcup_{k=n}^{\infty} A_k}(x)$$
$$= \min_{n\geq 1}\max_{k\geq n} I_{A_k}(x) = \inf_{n\geq 1}\sup_{k\geq n} I_{A_k}(x) = \overline{\lim_{n\to\infty}} I_{A_n}(x),$$
$$I_{\underline{\lim\limits_{n\to\infty}} A_n}(x) = I_{\bigcup_{n=1}^{\infty}\bigcap_{k\geq n} A_k}(x) = \max_{n\geq 1}\min_{k\geq n} I_{A_k}(x)$$
$$= \sup_{n\geq 1}\inf_{k\geq n} I_{A_k}(x) = \underline{\lim_{n\to\infty}} I_{A_n}(x). \qquad \square$$

习题

1. 证明： $(A\cup B)\setminus B = A \Longleftrightarrow A\cap B = \varnothing$.

2. 证明： $(A\setminus B)\cup B = A \Longleftrightarrow B\subset A$.

3. $(A\setminus B)\cup C = A\setminus(B\setminus C)$ 成立的充分必要条件是什么？

4. 证明下述等式：

1) $A\cap B = A\setminus(A\setminus B)$;

2) $A\setminus(B\setminus C) = (A\setminus B)\cup(A\cap C)$;

3) $A\setminus(B\cap C) = (A\setminus B)\cup(A\setminus C)$;

4) $A\setminus(B\cup C) = (A\setminus B)\cap(A\setminus C)$;

5) $(A\setminus B)\cap(C\setminus D) = (A\cap C)\setminus(B\cup D)$;

6) $(A\Delta B)\Delta C = A\Delta(B\Delta C)$;

7) $(A\Delta B)\cap C = (A\cap C)\Delta(B\cap C)$;

8) $B\cap(\bigcup_{\alpha\in I} A_\alpha) = \bigcup_{\alpha\in I}(B\cap A_\alpha)$;

9) $B\cup(\bigcap_{\alpha\in I} A_\alpha) = \bigcap_{\alpha\in I}(B\cup A_\alpha)$.

5. 下列等式是否成立？若不成立，有怎样的包含关系？

1) $A\cup(B\setminus C) = (A\cup B)\setminus(A\cup C)$;

2) $A\cup(B\Delta C)=(A\cup B)\Delta(A\cup C)$;

3) $A\setminus(B\cup C)=(A\setminus B)\cup(A\setminus C)$;

4) $(A\setminus B)\cup C=(A\cup C)\setminus B$.

6. 试化简集合 $(A\cup B^c\cup C^c)\cap(A\cup(B\cup C^c))$.

7. 设 $\{A_n: n=1,2,\cdots\}$ 为单调减集序列，则有

$$A_1=\Big(\bigcap_{n=1}^{\infty}A_n\Big)\bigcup\Big(\bigcup_{n=1}^{\infty}(A_n-A_{n+1})\Big),$$

且右端各项互不相交.

8. 设 $\mathfrak{R}$ 为 Ω 的一切子集组成的集类，则 $\mathfrak{R}$ 对集合的交 (看成乘法)、对称差 (看成加法) 运算作成一个环. Ω 是单位元，$\emptyset$ 是零元.

9. 设 $\{A_n: n=1,2,\cdots\}$ 是一集序列，令 $B_1=A_1$, $B_n=A_n\setminus(\bigcup_{k=1}^{n-1}A_k)$, 则 $B_n, n=1,2,\cdots$ 两两不交，且

$$\bigcup_{n=1}^{\infty}A_n=\bigcup_{n=1}^{\infty}B_n.$$

10. 试证明定理 7.

11. 试举一 $\{A_n: n=1,2,\cdots\}$ 不单调，但 $\lim\limits_{n\to\infty}A_n$ 存在的例.

12. 设

$$\begin{aligned}A_{2k+1}&=[0,2-\frac{1}{2k+1}],\\ A_{2k}&=[0,1+\frac{1}{2k}],\end{aligned}\qquad k=1,2,\cdots.$$

试求 $\underline{\lim}\limits_{n\to\infty}A_n$, 和 $\overline{\lim}\limits_{n\to\infty}A_n$.

13. 给定自然数 m 及 m 个集合 $B_0,\cdots,B_{m-1}$, 设 $A_n=B_k$, 当 m 整除 $n-k$ 时，试求 $\underline{\lim}\limits_{n\to\infty}A_n$, $\overline{\lim}\limits_{n\to\infty}A_n$.

14. 试证定理 10 的 1) 、 2) 、 3).

15. 设 $A, B \subset \Omega$, 试将 $I_{A \setminus B}, I_{A^c}, I_{A \Delta B}$ 用 I_A, I_B 表示出来.

§2. 映射与势

1. 定义 映射 $f: A \mapsto B$ 是指对一切 $x \in A$, 存在唯一的 $y \in B$, 使得 $f(x) = y$. 称 A 为 f 的 **定义域**; 对 $F \subset A$, 称 $f(F) := \{ y \in B : 存在\ x \in F, 使得\ f(x) = y \}$ 为 F 在 f 之下的 **象**, 而称 $f(A)$ 为 f 的 **值域**(显然应有 $f(A) \subset B$). 若有 $f(A) = B$, 则称 f 是 **满射**; 若由 $f(x_1) = f(x_2)$ 必得 $x_1 = x_2$, 则称 f 是 **单射**; 若 f 既是单射又是满射, 则称 f 是 A 到 B 上的 **一一映射(双射)**.

设 f 是 A 到 B 的映射， g 是 B 到 C 的映射，则把由等式

$$g \circ f(x) := g(f(x)),\ \forall\, x \in A,$$

定义的映射 $g \circ f: A \mapsto C$ 称为 g, f 的 **复合映射**.

注意： $g \circ f$ 与 $f \circ g$ 是不同的，且在上述情形 $f \circ g$ 未必有定义.

设 $f: A \mapsto B$, 则

$$f^{-1}(\{y\}) := \{ x \in A :\ f(x) = y \},\ y \in B;$$

$$f^{-1}(F) := \{ x \in A :\ f(x) \in F \},\ F \subset B.$$

且称 $f^{-1}(F)$ 为 F 在 f 之下的 **逆象**.

下面给出象与逆象的一些性质：设 $A_0, A_1, A_2 \subset A, B_1, B_2 \subset B$.

1) $A_1 \subset A_2 \Rightarrow f(A_1) \subset f(A_2)$;

2) $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$;

3) $f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$;

4) $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$;

5) $f^{-1}(B_1\cap B_2)=f^{-1}(B_1)\cap f^{-1}(B_2)$;

6) $f^{-1}(f(A_0))\supset A_0$.

若 f 为 A 到 B 的单射，则可定义 $f(A)(\subset B)$ 到 A 上的**逆映射** $f^{-1}:f^{-1}(y):=x$, 若 $f(x)=y$, 对任何 $y\in f(A)$.

注意此时 $f^{-1}(y)$ 与上面的 $f^{-1}(\{y\})$ 的意义不同.

2. 定义 给定集合 A,B 若存在 A 到 B 的一一映射，则称 A 与 B **对等**, 记作 $A\sim B$.

例1 $\mathbb{N}\sim 2\mathbb{N}:=\{\,2n:n\in\mathbb{N}\,\}$.

对等关系 "~" 满足：

(i) 自反性：对任何集合 A 有 $A\sim A$;

(ii) 传递性：若 $A\sim B, B\sim C$, 则 $A\sim C$;

(iii) 对称性：若 $A\sim B$, 则 $B\sim A$.

因此对等关系是一个等价关系. 按对等关系组成的集类称为此类中的集合的**势**. 对等集组成的类中集合的势相同. 集合 A 的势记作 $\bar{\bar{A}}$. 设 $n\in\mathbb{N}$, 则具有 n 个元素的集合作成的类是一个等价类，这类集合的共同标志是有 n 个元素，因此可以认为此类集合的势为 n. 对于无限集的情形则不同，请看下例.

例2 $(-1,1)\sim\mathbb{R}$. 取映射 $f(x)=\frac{x}{1-x^2}, x\in(-1,1)$. 则 f 是 (-1, 1) 到 $\mathbb{R}$ 的一一映射.

例3 (-1, 1]~(-1, 1).

注意到

$$(-1,1]\setminus(\{\,1-\frac{1}{n}:\,n\in\mathbb{N}\,\}\cup\{\,1\,\})=(-1,1)\setminus\{\,1-\frac{1}{n}:\,n\in\mathbb{N}\,\},$$

而

$$\{\,1-\frac{1}{n}:\,n\in\mathbb{N}\,\}\cup\{\,1\,\}\sim\{\,1-\frac{1}{n}:\,n\in\mathbb{N}\,\}.$$

这是因为可取一一映射

$$f: f(1)=0,\ f(1-\frac{1}{n})=1-\frac{1}{n+1}, n\geq 1.$$

实际上，无限集能和它的真子集对等，而有限集则不行. 有些集合论的专著就是利用这一事实定义有限集和无限集，再用自然数表示有限集的势.

由上述例 2 可以看出：有时很容易从直观上看出是对等的集合，如 (-1, 1] 和 (-1, 1)，但是证明其对等却并不容易. 为此我们证明下面的定理.

3. 定理 (F. Bernstein). 给定集合 A, B. 若存在 $A_1\subset A$, $B_1\subset B$. 且 $A_1\sim B$, $B_1\sim A$. 则 $A\sim B$.

证明 (i) 首先给出一个事实 (证明留作习题): 设 $\{A_\alpha:\alpha\in I\}$, $\{B_\alpha:\alpha\in I\}$ 是两个集类. 且对一切 $\alpha\in I$, $A_\alpha\sim B_\alpha$, $A_\alpha, \alpha\in I$ 两两不交, $B_\alpha, \alpha\in I$ 两两不交, 则

$$\bigcup_{\alpha\in I}A_\alpha\sim\bigcup_{\alpha\in I}B_\alpha.$$

(ii) 设 φ_1 是 A 到 B_1 的一一映射， φ_2 是 B 到 A_1 的一一映射. 令 $A_2:=\varphi_2(B_1)$, 则 $A\xrightarrow{\varphi_1}B_1\xrightarrow{\varphi_2}A_2$. 由 $B_1\subset B$ 知 $A_2=\varphi_2(B_1)\subset\varphi_2(B)=A_1$, 且 φ_2 也是 B_1 到 A_2 的一一映射, 即 $B_1\sim A_2$, 所以有

$$A\sim A_2,\quad 且\ A_2\subset A_1.$$

令 $B_2:=\varphi_1(A_1), A_3:=\varphi_2(B_2)$. 由 $B_2\subset\varphi_1(A)=B_1$ 知 $A_3\subset\varphi_2(B_1)=A_2$. 由 φ_1 是 A_1 到 B_2 的一一映射， φ_2 是 B_2 到 A_3 的一一映射，有 $A_1\sim B_2, B_2\sim A_3$. 从而

$$A_1\sim A_3,\ A_3\subset A_2.$$

类似地令

$$B_n:=\varphi_1(A_{n-1}), A_{n+1}:=\varphi_2(B_n), n\geq 3,$$

得 $\{A_n : n \in \mathbb{N}\}$, 满足

$$A \supset A_1 \supset A_2 \supset \cdots,$$

且在同一映射 $\varphi = \varphi_2 \circ \varphi_1$ 下有

$$A \sim A_2 \sim A_4 \sim \cdots,$$

$$A_1 \sim A_3 \sim A_5 \sim \cdots.$$

而

$$(1) \quad \begin{aligned} A &= \Big(\bigcap_{n=1}^{\infty} A_n\Big) \bigcup (A \setminus A_1) \bigcup (A_1 \setminus A_2) \bigcup (A_2 \setminus A_3) \bigcup \cdots \\ A_1 &= \Big(\bigcap_{n=1}^{\infty} A_n\Big) \bigcup (A_1 \setminus A_2) \bigcup (A_2 \setminus A_3) \bigcup (A_3 \setminus A_4) \bigcup \cdots, \end{aligned}$$

上述两式右端都是互不相交的集合之并. 注意到在同一映射 φ 下有

$$(2) \quad \begin{aligned} & A \setminus A_1 \sim A_2 \setminus A_3, \\ & A_2 \setminus A_3 \sim A_4 \setminus A_5 \\ & \cdots\cdots\cdots\cdots, \\ & A_{2n} \setminus A_{2n+1} \sim A_{2(n+1)} \setminus A_{2(n+1)+1}, \\ & \cdots\cdots\cdots\cdots, \end{aligned}$$

(为什么?) 利用 (i) 中的事实及 (1) (2), 即得 $A \sim A_1$. 已知 $A_1 \sim B$, 故得 $A \sim B$. □

4.推论 若 $A \subset B \subset C$ 且 $A \sim C$, 则 $A \sim B \sim C$.

证明 由 $C \supset B$、$B \supset A$ 及 $C \sim A$、$B \sim B$, 应用定理 3 即得 $C \sim B$.

5. 例 用 $\mathbb{R}$ 表示全体实数组成的集合, 则 $\mathbb{R}$ 中的任何有限区间及无限区间都与 $\mathbb{R}$ 对等, 即对任何 $a, b \in \mathbb{R}$, $a < b$,

$$[a, b] \sim (a, b] \sim [a, b) \sim (a, b) \sim (-\infty, b) \sim (a, \infty) \sim \mathbb{R}.$$

证明 我们知道 $(a,b)\sim(-1,1)\sim\mathbb{R}$, 而

$$(a,b)\subset(a,b]\subset\mathbb{R},\quad (a,b)\subset[a,b)\subset\mathbb{R},$$

$$(a,b)\subset[a,b]\subset\mathbb{R},\quad (a,b)\subset(a,\infty)\subset\mathbb{R},$$

$$(a,b)\subset(-\infty,b)\subset\mathbb{R}.$$

由此及推论知 $\mathbb{R}$ 中的任何有限区间及无限区间与 $\mathbb{R}$ 对等. □

注 对于 $(a,b)\sim(a,b]$ 也可象前面证明 $(-1,1)\sim(-1,1]$ 一样的方法证明. 事实上, 只要在 (a,b) 中取出一个集合 $\{a_n, n\in\mathbb{N}\}$. 例如, 令

$$a_n=a+\left(\frac{1}{b-a}+n\right)^{-1},\qquad n=1,2,\cdots,$$

则

$$(a,b)\setminus\{a_n,n\in\mathbb{N}\}=(a,b]\setminus(\{a_n,n\in\mathbb{N}\}\cup\{b\}),$$

因而对等. 而

$$\{a_n,n\in\mathbb{N}\}\sim\{b,a_n,n\in\mathbb{N}\},$$

即

$$a_1\longleftrightarrow b;\qquad a_n\longleftrightarrow a_{n-1},\quad n\geq 2.$$

于是 $(a,b)\sim(a,b]$.

习题

1. 证明定理 3 的证明 (i) 中所提出的事实.

2. 试作开上半平面与开单位圆间的一一映射.

3. 设集合 A 有 n 个元素 $(n=1,2,\cdots)$, 在 A 的子集和它的余集 (对 A) 间建立一一对应, 由此证明

$$C_n^k=C_n^{n-k},\qquad 0\leq k\leq n.$$

其中 C_n^k 表示从 n 个元中取出 k 个元的组合数.

§3. 可数集

1. 定义 若 $A \sim \mathbb{N}$, 则称 A 为 **可数集 (或可列集)**. 可数集 $\mathbb{N}$ 的势记作 $\aleph_0$, 读作 "阿列夫零". 不是可数集的无限集称为 **不可数集**.

本节和下节讨论可数集与不可数集，特别要指出一些常见集哪些是可数集, 哪些是不可数集. 这是很有用的, 因为可数与不可数情形常常有本质的不同.

例1 偶数集，奇数集，整数集都是可数集.

2. 定理 任何无限集必含有可数子集.

证明 设 A 为无限集，从 A 中取出一元素记为 a_1. 因为 A 是无限集，所以 $A \setminus \{ a_1 \}$ 非空，从中取出一元素记为 a_2, 显然 $a_2 \neq a_1$. 设已从 A 中取出互不相同的元素 $a_1, a_2, \cdots, a_n$, 因为 A 是无限集，$A \setminus \{ a_1, a_2, \cdots, a_n \}$ 非空，还可以从中取出元素 a_{n+1}. 显然它不同于 $a_1, a_2, \cdots, a_n$. 由归纳法，得到了 A 中互不相同的元素组成的集合 $A_1 = \{ a_1, \cdots, a_n, \cdots \} \subset A$. 显然 A_1 是 A 的可数子集. □

3. 定理 1) 可数集的子集至多可数；

2) 若 A 是可数集，B 是有限集，则 $A \cup B$ 为可数集.

证明 1) 设 A 为可数集. 将 A 中元素编号如下：

$$a_1, \cdots, a_n, \cdots.$$

设 B 是 A 的任一非空子集 (若 $B = \varnothing$, 结论显然成立). 则 B 中元素是上述序列的一个子列，即

$$B = \{ a_{n_1}, \cdots, a_{n_k}, \cdots \}.$$

足码 $n_1, n_2, \cdots, n_k, \cdots$, 中如有最大者，则 B 是有限集，否则 B 为无限集. 让 a_k 与 k 对应，即知 B 为可数集.

2) 记 $A=\{\,a_1,\cdots,a_n,\cdots\,\}$, $B\setminus A=\{\,b_1,b_2,\cdots,b_m\,\}$ (设 $B\setminus A$ 有 m 个元素), 则 $A\cup B=\{\,b_1,b_2,\cdots,b_m,a_1,\cdots,a_n,\cdots\,\}$ 也是可数集. □

例1　$\mathbb{Z}^2=\{\,(n,m):\ n,m\in\mathbb{Z}\,\}$ 和 Q(有理数集) 都是可数集.

证明　下面定义的 φ:

$$\varphi(0):=0,\varphi(2n-1):=n,\varphi(2n)=-n,\quad n\geq 1$$

是 $\mathbb{Z}_+$ 到 $\mathbb{Z}$ 的一一映射 (为什么？), 因而如下定义的

$$\tilde{\varphi}:\quad \tilde{\varphi}(m,n):=(\varphi(m),\varphi(n)),\quad (m,n)\in\mathbb{Z}_+^2$$

为 $\mathbb{Z}_+^2$ 到 $\mathbb{Z}^2$ 的一一映射, 所以有 $\mathbb{Z}^2\sim\mathbb{Z}_+^2$.

若将 $\mathbb{Z}_+^2$ 的元按如下方式排列:

$$\begin{array}{ccccccccc}
(0,0) & \rightarrow & (0,1) & & (0,2) & \rightarrow & (0,3) & \rightarrow & \cdots \\
 & \swarrow & & \nearrow & & \swarrow & & \nearrow & \\
(1,0) & & (1,1) & & (1,2) & & (1,3) & & \cdots \\
\downarrow & \nearrow & & \swarrow & & \nearrow & & \swarrow & \\
(2,0) & & (2,1) & & (2,2) & & (2,3) & & \cdots \\
 & \swarrow & & \nearrow & & \swarrow & & \nearrow & \\
(3,0) & & (3,1) & & (3,2) & & (3,3) & & \cdots \\
\downarrow & \nearrow & & \swarrow & & \nearrow & & \swarrow & \\
(4,0) & & (4,1) & & (4,2) & & (4,3) & & \cdots \\
 & \swarrow & & \nearrow & & \swarrow & & & \\
\vdots & & \vdots & & \vdots & & \vdots & & \ddots
\end{array}$$

则可知 $\mathbb{Z}_+^2\sim\mathbb{N}$. 故有 $\mathbb{Z}^2\sim\mathbb{N}$.

令

$$Q:=\{\,\frac{p}{q}:q,p\in\mathbb{Z},(p,q)=1,q>0\,\};$$

$$\widetilde{Q}=\{\,(p,q):q,p\in\mathbb{Z},(p,q)=1,q>0\,\},$$

其中 $(p,q)=1$ 表示 p,q 的互素 (即最大公因数为 1). 则 $Q\sim\widetilde{Q}$. 而 $\widetilde{Q}\subset\mathbb{Z}^2$. 由定理 3 及 (i) 知 $\widetilde{Q}$ 至多可数, 而 $\widetilde{Q}$ 为无限集 (为什么?). 所以 $\widetilde{Q}$ 为可数集. 因而 Q 为可数集. □

4. 定理 若 $A_i(i=1,2,\cdots)$ 是可数集, 则 $\bigcup_{i=1}^n A_i$, $n\in\mathbb{N}$ 及 $\bigcup_{i=1}^\infty A_i$ 都是可数集.

因而若 $A_i(i=1,2,\cdots)$ 至多可数 (有限), 则 $\bigcup_{i=1}^n A_i$, $n\in\mathbb{N}$, 至多可数 (有限); $\bigcup_{i=1}^\infty A_i$ 至多可数.

证明 由于后一结论是前一结论的直接推论或为显然事实, 只需证前一结论. 又由于证法类似, 所以只证后者. 因为 $A_i(i=1,2,\cdots)$ 是可数集, 可设 $A_i:=\{\,a_{ik}:\ k=1,2,\cdots\}$. 于是

$$\bigcup_{i=1}^\infty A_i=\{\,a_{ik}:\ (i,k)\in\mathbb{N}^2\,\}.$$

由于 A_i, $i\in\mathbb{N}$ 可能相交, a_{ik}, $(i,k)\in\mathbb{N}^2$ 中有重复元, 因而 $\bigcup_{i=1}^\infty A_i$ 与 $\mathbb{N}^2\subset\mathbb{Z}^2$ 的一个子集 B 对等. 由定理 3 和例 1 知 B, 因而 $\bigcup_{i=1}^\infty A_i$, 至多可数. 但它已包含了一个可数集 A_1, 不会是有限集, 所以 $\bigcup_{i=1}^\infty A_i$ 是可数集. □

5. 定理 若 $A_i(i=1,2,\cdots)$ 是可数集, 则对任何 $n\in\mathbb{N}$,

$$A_1\times A_2\times\cdots\times A_n:=\{\,(a_1,a_2,\cdots,a_n):\ a_i\in A_i,\ i=1,2,\cdots,n\,\}$$

是可数集. 此集称之为 $A_1,A_2,\cdots,A_n$ 的 **直乘积**. 当 $A_1=\cdots=A_n=A$ 时, 记 $A_1\times A_2\times\cdots\times A_n=A^n$.

证明 用数学归纳法证明. $n=1$ 时, 定理显然成立.

1) 当 $n=2$ 时, 记 $A_i=\{\,a_{ik}:\ k\in\mathbb{Z}_+\,\}$, $i=1,2$, 则

$$A_1\times A_2=\{\,(a_{1k},a_{2j}):\ (k,j)\in\mathbb{Z}_+^2\,\}$$

与 $\mathbb{Z}_+^2$ 对等 (因为 $(a_{1k},a_{2j})\mapsto(k,j)$ 是它们之间的一一映射). 所以 $A_1\times A_2$ 是可数集.

2) 设 $n=k$ 时定理成立，往证 $n=k+1$ 时定理也成立.

注意到 $(A_1\times\cdots\times A_k)\times A_{k+1}\sim A_1\times\cdots\times A_k\times A_{k+1}$，由归纳假设，$A_1\times\cdots\times A_k$ 可数，则由 (i) 知 $(A_1\times\cdots\times A_k)\times A_{k+1}$ 可数，从而有 $A_1\times\cdots\times A_k\times A_{k+1}$ 可数，故对任何 $n\in\mathbb{N}$, $A_1\times\cdots\times A_n$ 可数.　　□

6. 例1　全体整系数多项式组成的集为可数集合.

证明　由于两多项式相等当且仅当次数相同且对应幂的系数相等，

$$\begin{aligned}
&\Big\{\sum_{i=0}^{n}a_ix^i:\ a_i\in\mathbb{Z}, i=0,1,2,\cdots,n, n\in\mathbb{N}\Big\}\\
&\sim\{(a_0,a_1,\cdots,a_n):\ a_i\in\mathbb{Z}, i=0,1,2,\cdots,n, n\in\mathbb{N}\}\\
&=\bigcup_{n\in\mathbb{N}}\{(a_0,\cdots,a_n):\ a_i\in\mathbb{Z},\ i=0,1,2,\cdots,n\}=\bigcup_{n\in\mathbb{N}}\mathbb{Z}^n.
\end{aligned}$$

由定理 5 知 $\mathbb{Z}^n$ 是可数集，再由定理 4 知 $\bigcup_{n\in\mathbb{N}}\mathbb{Z}^m$ 是可数集.

例2　整系数多项式的根称为 **代数数**. 全体代数数的集是可数集 (证明留给读者，作为练习).

另外，不是代数数的实 (复) 数称为 **超越数**. 我们在以后将会知道，全体实数组成的集合不可数，所以超越数集不可数. 但能具体指出的超越数却很少. 现在知道 e,π 以及 α^β, (α 是不为 0, 1 的代数数，β 是非有理数的代数数)，是超越数.

例3　设 $f(x)$ 是直线上的单调函数，则它的全体间断点组成的集至多可数.

证明　不防设 f 单调增，对任何 $x\in\mathbb{R}$,

$$f(x-0):=\lim_{y\to x-0}f(y),\ f(x+0):=\lim_{y\to x+0}f(y),$$

则 x_0 是 $f(x)$ 的间断点当且仅当

$$f(x_0+0)-f(x_0-0)>0,$$

且若 x_1 为异于 x_0 的间断点，则

$$(f(x_0-0), f(x_0+0)) \cap (f(x_1-0), f(x_1+0)) = \varnothing,$$

再用本节习题 2 的结果可证： f 的间断点至多可数.

注意: 可数集的元素可用自然数编号，不一定能按大小次序排列.

习题

1. $\mathbb{R}^d$ 中以有理点为中心， 以正有理数为半径的球的全体是可数集.

2. 直线上一个由长度不为零的互不相交的开区间组成的集至多可数.

3. 证明任一可数集的所有有限子集的全体组成可数集.

4. 给定平面上一个集，若此集中的任意两点间距离大于某个固定正数 α, 则此集至多是可数集.

5. 设 A 是有限集或可数集， B 是无限集，则 $A \cup B \sim B$.

§4. 不可数集

1. 定理 [0,1] 是不可数集. [0,1] 的势记作 $\aleph$, 读作"阿列夫".

证明 用反证法. 设 [0, 1] 可数，于是有

$$[0,1] = \{x_1, x_2, \cdots, x_n, \cdots\} \tag{1}$$

显然 $[0, \frac{1}{3}]$, $[\frac{1}{3}, \frac{2}{3}]$, $[\frac{2}{3}, 1]$ 中至少有一个不含 x_1, 用 I_1 表示此闭区间，再将 I_1 等分成三个闭区间，其中至少有一个不含 x_2, 用 I_2 表示此闭区间，且有 $I_1 \supset I_2$; 如此类推下去，由归纳定义可得一闭区间套:

$$\{I_n : n \in \mathbb{N}\}: \ I_1 \supset I_2 \supset \cdots \supset I_n \supset \cdots;$$

$$x_n \notin I_n \ (n \in \mathbb{N}); \ |I_n| = \frac{1}{3^n} \to 0 (n \to \infty).$$

由区间套引理知有一 $\zeta \in I_n$ (对任何 $n \in \mathbb{N}$). 显然 $\zeta \in [0,1]$, 但对任何 $n \in \mathbb{N}, \zeta \neq x_n$. 这与 (1) 矛盾. □

推论 直线上任一非退化区间的势都是 $\aleph$.

2. 定理 全体实数列作成的集合 (记作 E^∞) 的势为 $\aleph(=\aleph^{\aleph_0})$.

证明 记

$$E^\infty = \{\, (x_1, \cdots, x_n, \cdots) : \ x_i \in \mathbb{R}, i \in \mathbb{N} \,\},$$

$$B = \{\, (x_1, \cdots, x_n, \cdots) : \ x_i \in (0,1), i \in \mathbb{N} \,\}.$$

定义映射 $\varphi : B \mapsto E^\infty$ 为：对任何 $x \in B$,

$$\varphi(x) := \left(\tan(x_1 - \frac{1}{2})\pi, \cdots, \tan(x_n - \frac{1}{2})\pi, \cdots\right),$$

则 φ 是 B 到 E^∞ 的一一映射，所以 $E^\infty \sim B$, 只需证 $\bar{\bar{B}} = \aleph$. 为此往证： $B \sim (0,1)$.

首先，将 $(0,1)$ 中的任一 x 与 B 中的点 $\bar{x} = (x, x, \cdots, x, \cdots)$ 对应，故有

$$(0,1) \sim B_1 := \{\, \bar{x} = (x, x, \cdots, x, \cdots) : \ x \in (0,1) \,\} \subset B.$$

其次，用十进制无限小数表示 (此时表示方法唯一) B 中任一 x 的分量：

$$x_1 = 0.x_{11}x_{12}\cdots x_{1n}\cdots$$

$$x_2 = 0.x_{21}x_{22}\cdots x_{2n}\cdots$$

$$\cdots$$

$$x_n = 0.x_{n1}x_{n2}\cdots x_{nn}\cdots$$

$$\cdots$$

现在依照

$$\psi(x) := 0.x_{11}x_{12}x_{21}\cdots x_{1n}x_{2(n-1)}\cdots x_{n1}\cdots$$

将 x 映入 $(0,1)$, 于是 $\psi(x)\in(0,1)$, 且当 $x\neq y$, $x,y\in B$ 时, $\psi(x)\neq\psi(y)$. 故 $(0,1)\supset\psi(B)\sim B$. 由 Bernstein 定理知

$$(0,1)\sim B\sim E^{\infty}.\qquad\square$$

3. 定理 全体无限二进位小数组成的集合是 $(0,1]$. 因而

$$X:=\{\,(x_1,x_2,\cdots):\ x_n\in\{\,0,1\,\},n\in\mathbb{N}\,\}$$

的势为 $\aleph$.

证明 $(0,1]$ 中的实数 x 若不具有 $\frac{m}{2^n}$, $m,n\in\mathbb{N}$, 的形式, 则必能唯一地表成无限二进位小数:

$$x=\frac{x_1}{2}+\frac{x_2}{2^2}+\cdots+\frac{x_n}{2^n}+\cdots$$

其中 $x_n\in\{\,0,1\,\},n\in\mathbb{N}$, 且有无限多个 $x_n=1$. 若 $x=\frac{m}{2^n}$, 对某对 $n,m\in\mathbb{N}$ 成立 (不妨设 m 为奇数), 则有且仅有两种表示:

$$x=\frac{x_1}{2}+\cdots+\frac{x_{n-1}}{2^{n-1}}+\frac{1}{2^n}$$

$$x=\frac{x_1}{2}+\cdots+\frac{x_{n-1}}{2^{n-1}}+\frac{0}{2^n}+\sum_{k=n+1}^{\infty}\frac{1}{2^k}.$$

前者为有限二进位小数. 故综合上面讨论即得第一结论.

由上述结论知全体无限二进位小数的势为 $\aleph$, 且与

$$\tilde{X}:=\{(x_1,\cdots,x_n,\cdots):x_n\in\{0,1\},n\in\mathbb{N},\text{且其中有无限多个 }1\}$$

对等, 即 $X\supset\tilde{X}$, 而 $\tilde{X}$ 的势为 $\aleph$. 由定理 2 及推论 2.4 知 $X\sim\tilde{X}$, 因而 X 的势为 $\aleph$. $\square$

定理 3 中的 X 有一种物理解释: 设想在 $\mathbb{N}$ (甚至是 $\mathbb{Z}^d$) 的每一点上有一粒子, 而每一粒子具有两种状态: 0 和 1, 则整个粒子系统的状态就可以用 X 中的元素表示.

用定理 2, 3 可以证明一些集合的势为 $\aleph$. 我们在习题中列举了若干, 让读者证明. 其中一个例子是可数集的全体子集组成的集合的势是 $\aleph$, 它比可数集的势要 "大". 这个事实具有一般性. 为此先给出

4. 定义 设 $\bar{\bar{A}} = \alpha, \bar{\bar{B}} = \beta$, 若 A 不与 B 对等而与 B 的一个子集对等，则称 A 的势小于 B 的势，记作 $\alpha < \beta$ (或 $\beta > \alpha$).

5. 定理 $\bar{\bar{A}} = \mu$, 用 2^μ 表示 A 的全体子集组成的集合的势，则 $\mu < 2^\mu$.

证明 用 $\mathcal{P}(A)$ 表示 A 的全体子集组成的集合，记

$$\widetilde{A} := \{\, \{x\} : x \in A \,\},$$

则 $A \sim \widetilde{A} \subset \mathcal{P}(A)$. 于是只需证： $A \nsim \mathcal{P}(A)$.

假设 $A \sim \mathcal{P}(A)$, 令 f 为由 A 到 $P(A)$ 的一一映射，造集合

$$B := \{\, x \in A : x \notin f(x) \,\} \subset A,$$

于是 $B \in \mathcal{P}(A)$, 因而存在 $y \in A$ 使 $f(y) = B$. 现在分析 y 与 B 的关系：

1) 若 $y \in B$, 则由 B 的定义 $y \notin f(y) = B$, 得出矛盾；

2) 若 $y \notin B$, 仍由 B 的定义， $y \in f(y) = B$, 也得出矛盾.

所以无论怎样都引出矛盾. 因此 A 不与 $\mathcal{P}(A)$ 对等. 故由定义 4 知 $\mu < 2^\mu$. □

关于集合的势的理论 (简称势论) 很丰富和深刻，我们就不再讨论了.

习题

1. 证明：定义在 $[a, b]$ 上的连续函数的全体组成的集 $C[a, b]$ 的势为 $\aleph$.

2. (1) 证明定义在 $[a, b]$ 上的右连续的单调函数全体的势为 $\aleph$.
   (2) 定义在 $[a, b]$ 上的单调函数全体具有什么样的势？

3. 证明自然数列全体的势为 $\aleph$.

4. 证明自然数的全体子集组成的集 $\mathcal{P}(\mathrm{N})$ 的势为 $\aleph$.

第二章 距离空间

§1. 定义及例

1. 引言 极限运算是分析数学中最重要的运算，而在数学分析（无论一元，还是多元）中，极限概念里起着基础作用的是点的距离. 实际上，一系列基本事实只与距离有关. 这种情况自然引导出距离的推广而给出“距离空间”的概念，它是实数轴 $\mathbb{R}$ 及 d 维实空间 $\mathbb{R}^d$ 的推广.

这一章有点承先启后的作用，它承继 $\mathbb{R}^d$ 的距离，又是拓扑空间的特殊情形，所以要求大家学习时要特别注意两点：

1) 通过实直线 $\mathbb{R}$ 以及 $\mathbb{R}^2$、$\mathbb{R}^3$、$\cdots,\mathbb{R}^d$ 来理解它，很多结论都一眼可以看出；

2) 要能丢掉拐杖（它的确与 $\mathbb{R}^d$ 有本质的不同），从距离空间的概念出发进行论证. 这样今后才能学好拓扑学（也包括概率理论）.

2. 定义 设 E 为一集，映射 $d: E\times E\mapsto\mathbb{R}$, 且满足：

Ⅰ) 对一切 $x,y\in E,\quad d(x,y)\geq 0$;

Ⅱ) $d(x,y)=0 \iff x=y$;

Ⅲ) 对一切 $x,y\in E,\quad d(x,y)=d(y,x)$;

Ⅳ) 对一切 $x,y,z\in E,\quad d(x,y)\leq d(x,z)+d(z,y)$

称 d 为 E 上的一个 **距离**，称 (E,d)（或简单写成 E）为（以 d 为距离的）**距离空间**. 本节中常用 $d,d_1,\cdots,\rho,\rho_1,\cdots$ 表示距离.

例如，对一切 $x, y \in \mathbb{R}$, $d(x,y) := |x-y|$, 则 d 即为 $\mathbb{R}$ 的距离.

3. **例1** 给定 E, 定义

$$d(x,y) := 1 - \delta(x,y) = \begin{cases} 0, & x = y; \\ 1, & x \neq y; \end{cases} \quad x, y \in E.$$

称 (E,d) 为 **离散距离空间**.

例2 $\mathbb{R}^d = \{x : x = (x_1, \cdots, x_d) : x_k \in \mathbb{R},\ k = 1, \cdots, d\}$, 其中 $d \in \mathbb{N}$, 中的各种距离.

在 $\mathbb{R}^d$ 中除离散距离外，可以定义以下各种距离：对于 $p \geq 1$,

$$(1) \qquad d_p(x,y) := \left[\sum_{k=1}^{d} |x_k - y_k|^p\right]^{\frac{1}{p}}, \qquad x, y \in \mathbb{R}^d,$$

在 (1) 中当 $d = 1$, $p = 1$ 时即为直线上两点的距离，当 $p = 2$ 时即通常的欧氏距离 (在 $d = 2$, 3 时，即平面或空间中通常的两点间距离).

$$(2) \qquad d_\infty(x,y) := \max_{1 \leq k \leq d} |x_k - y_k|, \qquad x, y \in \mathbb{R}^d.$$

证明 显然 $d_\infty(x,y)$ 是 $\mathbb{R}^d$ 上的距离，不需再证. 还可以证明：对任何 $x, y \in \mathbb{R}^d$,

$$(3) \qquad \lim_{p \to \infty} d_p(x,y) = d_\infty(x,y)$$

(留作习题). d_1 是距离也是显然的；要证 d_p 是距离，主要是证明它满足三角不等式. 对于 d_2, 通常用 Schwarz 不等式证明. 在这里，我们用不等式

$$(4) \qquad a^{\frac{1}{p}} b^{\frac{1}{q}} \leq \frac{a}{p} + \frac{b}{q}, \qquad a, b \in \mathbb{R}_+,$$

其中 $p, q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, $\mathbb{R}_+$ 表示全体非负实数组成的集 (证明留作习题), 来证明一般情形.

首先证明有用的 **Hölder 不等式** 及 **Minkowski 不等式**. 设

$$\{a_1, \cdots, a_d, b_1, \cdots, b_d\} \subset \mathbb{R}_+, p > 1, q > 1, \frac{1}{p} + \frac{1}{q} = 1,$$

则

(5) Hölder 不等式: $\displaystyle\sum_{k=1}^{d} a_k b_k \leq \left[\sum_{k=1}^{d} a_k^p\right]^{\frac{1}{p}} \left[\sum_{k=1}^{d} b_k^q\right]^{\frac{1}{q}}$.

(6) Minkowski 不等式: $\displaystyle\left[\sum_{k=1}^{d} (a_k + b_k)^p\right]^{\frac{1}{p}} \leq \left[\sum_{k=1}^{d} a_k^p\right]^{\frac{1}{p}} + \left[\sum_{k=1}^{d} b_k^p\right]^{\frac{1}{p}}$.

当 $\{a_k\}$ 全为零, 或 $\{b_k\}$ 全为零时, 不等式自然成立. 故不妨设它们都不为零. 以 $\dfrac{a_k^p}{\sum_{i=1}^d a_i^p}$, $\dfrac{b_k^q}{\sum_{i=1}^d b_i^q}$ 分别代换 (4) 中的 a, b, 再对 $k = 1, 2, \cdots, d$, 求和即得 (5). 其次应用 (5) 得

$$\begin{aligned}\sum_{k=1}^{d} (a_k + b_k)^p &= \sum_{k=1}^{d} (a_k + b_k)^{p-1} a_k + \sum_{k=1}^{d} (a_k + b_k)^{p-1} b_k \\ &\leq \left[\sum_{k=1}^{d} (a_k + b_k)^{(p-1)q}\right]^{\frac{1}{q}} \left(\left[\sum_{k=1}^{d} a_k^p\right]^{\frac{1}{p}} + \left[\sum_{k=1}^{d} b_k^p\right]^{\frac{1}{p}}\right).\end{aligned}$$

将上式两边除以 $\left[\sum_{k=1}^d (a_k + b_k)^{(p-1)q}\right]^{\frac{1}{q}}$, 注意到 $\frac{1}{p} + \frac{1}{q} = 1$ 及 $(p-1)q = p$, 即得 (6) 式.

由 (6) 很容易证明 d_p 是 $\mathbb{R}^d$ 上的距离. 事实上, 若 $x = (x_1, \cdots, x_d)$, $y = (y_1, \cdots, y_d)$, 及 $z = (z_1, \cdots, z_d)$, 则

$$\begin{aligned}d_p(x, y) &= \left[\sum_{k=1}^{d} |x_k - y_k|^p\right]^{\frac{1}{p}} \leq \left[\sum_{k=1}^{d} \left(|x_k - y_k| + |z_k - y_k|\right)^p\right]^{\frac{1}{p}} \\ &\leq \left[\sum_{k=1}^{d} |x_k - z_k|^p\right]^{\frac{1}{p}} + \left[\sum_{k=1}^{d} |z_k - y_k|^p\right]^{\frac{1}{p}} \\ &= d_p(x, z) + d_p(z, y). \quad \square\end{aligned}$$

这个例子可以大大推广. 在习题中, 读者看到一些稍微一般的情形. 对于一般情形, 我们留待积分理论中讨论.

例3 分布函数的距离. 设 $\mathcal{D}_F$ 表示一切概率分布函数 (在本书中, 假设它们右连续) 组成的集. 对一切 $F, G \in \mathcal{D}_F$, 定义

$$L(F,G) := \inf\{\varepsilon > 0: \ F(x-\varepsilon) - \varepsilon \le G(x) \le F(x+\varepsilon) + \varepsilon, \ \forall\, x \in \mathbb{R}\}, \tag{7}$$

则 L 是 $\mathcal{D}_F$ 上的一个距离, 称为 Lèvy **距离**.

证明 显然 $L(F,G) \ge 0$, 其次, 由

$$F(x-\varepsilon) - \varepsilon \le G(x) \le F(x+\varepsilon) + \varepsilon, \qquad \forall\, x \in \mathbb{R} \tag{8}$$

即得

$$F(x-\varepsilon) \le G(x) + \varepsilon, \qquad G(x) - \varepsilon \le F(x+\varepsilon), \qquad \forall\, x \in \mathbb{R},$$

因而, 有

$$G(x-\varepsilon) - \varepsilon \le F(x) \le G(x+\varepsilon) + \varepsilon, \qquad \forall\, x \in \mathbb{R}. \tag{9}$$

反之, (9) $\Longrightarrow$ (8), 故得 $L(F,G) = L(G,F)$, 即 III) 成立.

再次, 若 $F = G$, 则 $\forall\, \varepsilon > 0$, (8) 成立, 故有 $L(F,G) = 0$, 反之, 若 $L(F,G) = 0$, 则存在 $\{\varepsilon_n\}$, $0 < \varepsilon_n \downarrow 0$, 使 (8) 对任何 $\varepsilon = \varepsilon_n$ 成立. 令 $n \to \infty$, 则由右边不等式及 F 右连续知

$$G(x) \le F(x), \qquad \forall\, x \in \mathbb{R}.$$

再由 $L(G,F) = L(F,G) = 0$ 及上述证法即得

$$F(x) \le G(x), \qquad \forall\, x \in \mathbb{R}.$$

故得 $F = G$, 即 II) 成立.

最后, 设 $F, G, H \in \mathcal{D}_F$, 往证 $L(F,G) \le L(F,H) + L(H,G)$. 若 $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, 使得

$$F(x-\varepsilon_1) - \varepsilon_1 \le H(x) \le F(x+\varepsilon_1) + \varepsilon_1, \qquad \forall\, x \in \mathbb{R},$$

$$H(x-\varepsilon_2) - \varepsilon_2 \le G(x) \le H(x+\varepsilon_2) + \varepsilon_2, \qquad \forall\, x \in \mathbb{R}.$$

都成立，则对任何 $x\in\mathbb{R}$，有

$$\begin{aligned}F(x-\varepsilon_2-\varepsilon_1)-\varepsilon_2-\varepsilon_1&=[F((x-\varepsilon_2)-\varepsilon_1)-\varepsilon_1]-\varepsilon_2\\&\le H(x-\varepsilon_2)-\varepsilon_2\le G(x)\\&\le H(x+\varepsilon_2)+\varepsilon_2\le F(x+\varepsilon_1+\varepsilon_2)+\varepsilon_1+\varepsilon_2.\end{aligned}$$

故有

$$L(F,G)\le\varepsilon_1+\varepsilon_2.$$

令 $\varepsilon_1\downarrow L(F,H)$，$\varepsilon_2\downarrow L(H,G)$，即得 IV) □

概率分布函数列 $\{F_n\}$ 弱收敛于概率分布函数 F 是指对任何 $x\in C(F)$，$F_n(x)\to F(x)$. 其中 $C(F)$ 表示 F 的连续点集. 我们证明下述

4. 定理 概率分布函数列 $\{F_n\}$ 弱收敛于概率分布函数 F 的充分必要条件是当 $n\to\infty$ 时，$L(F_n,F)\to 0$.

证明 先证明条件的充分性.

任给 $x_0\in C(F)$. 对任何 $\varepsilon>0$，存在 $\delta>0$，使得对任何 $x\in(x_0-\delta,x_0+\delta)$，有 $|F(x)-F(x_0)|<\dfrac{\varepsilon}{2}$. 由 $L(F_n,F)\to 0, n\to\infty$ 知，对于 $\varepsilon_1:=\delta\wedge\frac{\varepsilon}{2}$，存在 n_0，当 $n\geqslant n_0$ 时，有 $L(F_n,F)<\varepsilon_1$，即

$$\inf\{\alpha: F(x-\alpha)-\alpha<F_n(x)<F(x+\alpha)+\alpha,\forall\, x\in\mathbb{R}\}<\varepsilon_1.$$

于是，存在 $\alpha<\varepsilon_1$，使得对任何 $x\in\mathbb{R}$，有

$$F(x-\alpha)-\alpha<F_n(x)<F(x+\alpha)+\alpha.$$

特别，对于 x_0，有

$$F_n(x_0)>F(x_0-\alpha)-\alpha>F(x_0)-\frac{\varepsilon}{2}-\alpha>F(x_0)-\varepsilon,$$
$$F_n(x_0)<F(x_0+\alpha)+\alpha<F(x_0)+\frac{\varepsilon}{2}+\alpha<F(x_0)+\varepsilon,$$

即

$$|F_n(x_0)-F(x_0)|<\varepsilon.$$

条件的充分性获证.

再证明条件的必要性.

1). 任给 $\varepsilon>0$, 存在 x_0, 使得 $-x_0, x_0\in C(F)$, 且 $F(-x_0)\vee(1-F(x_0))<\frac{\varepsilon}{2}$, 因而, 存在 $n_1\in\mathbb{N}$, 使得当 $n\geqslant n_1$ 时, $F_n(-x_0)<F(-x_0)+\frac{\varepsilon}{2}<\varepsilon$ 与 $F_n(x_0)>F(x_0)-\frac{\varepsilon}{2}>1-\varepsilon$ 同时成立. 于是当 $x\in(-\infty,-x_0], n\geqslant n_1$ 时有

$$\begin{aligned}F(x-\varepsilon)-\varepsilon&\leqslant F(-x_0)-\varepsilon<0\leqslant F_n(x)\\&\leqslant F_n(-x_0)<\varepsilon\leqslant F(x+\varepsilon)+\varepsilon,\end{aligned}$$

而当 $x\in[x_0,+\infty), n\geqslant n_1$ 时有

$$\begin{aligned}F(x-\varepsilon)-\varepsilon&\leqslant 1-\varepsilon<F(x_0)-\frac{\varepsilon}{2}\\&<F_n(x)\leqslant 1<F(x_0)+\varepsilon\leqslant F(x+\varepsilon)+\varepsilon.\end{aligned}$$

2). 由于连续点的稠密性, 对任何 $y\in[-x_0,x_0]$, 存在 x_y, x_y' $\in C(F)$, 使得

$$y-\frac{\varepsilon}{2}<x_y<y<x_y'<y+\frac{\varepsilon}{2}.$$

于是

$$\bigcup_{y\in[-x_0,x_0]}(x_y,x_y')\supset[-x_0,x_0].$$

由有限覆盖定理, 存在 $\{y_1,y_2,\cdots,y_m\}\subset[-x_0-\frac{\varepsilon}{2},x_0+\frac{\varepsilon}{2}]$, 使得

$$\bigcup_{1\leqslant k\leqslant m}(x_{y_k},x_{y_k}')\supset[-x_0,x_0],$$

且

$$0<x_{y_k}'-x_{y_k}<\varepsilon,\qquad\forall\,1\leqslant k\leqslant m.$$

于是存在 n_2, 对一切 $n\geqslant n_2$, $1\leqslant k\leqslant m$, 有

$$F_n(x_{y_k})>F(x_{y_k})-\varepsilon,\quad F_n(x_{y_k}')<F(x_{y_k}')+\varepsilon.$$

从而对任何 $x \in [-x_0, x_0]$, 存在 $1 \leqslant k \leqslant m$, 使得

$$x \in (x_{y_k}, x'_{y_k}), \quad \text{且 } x - \varepsilon < x_{y_k} < x'_{y_k} < x + \varepsilon.$$

这样,

$$\begin{aligned} F(x-\varepsilon) - \varepsilon &\leqslant F(x_{y_k}) - \varepsilon < F_n(x_{y_k}) \\ &\leqslant F_n(x) \leqslant F_n(x'_{y_k}) < F(x'_{y_k}) + \varepsilon \leqslant F(x+\varepsilon) + \varepsilon. \end{aligned}$$

3). 综合 1), 2) 即得: 任给 $\varepsilon > 0$, 令 $n_0 := n_1 \vee n_2$, 则对任何 $n \geqslant n_0$, $L(F, F_n) \leqslant \varepsilon$. 这就完成了定理的证明. □

距离象 $\mathbb{R}$ 中的点 x, y 具有性质 $||x| - |y|| \leqslant |x - y|$ 一样, 有

5. 引理 设 (E, d) 为距离空间, 则对任何 $x, y, z \in E$, 有

$$|d(x, z) - d(y, z)| \leqslant d(x, y). \tag{1}$$

证明 因为

$$d(x, z) \leq d(x, y) + d(y, z) \implies d(x, z) - d(y, z) \leq d(x, y),$$

再由对称性得

$$d(y, z) - d(x, z) \leq d(x, y).$$

故

$$|d(x, z) - d(y, z)| \leq d(x, y). \quad \square$$

习题

1. 若 $p > 1$, $q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, 则对任何 $a, b \in \mathbb{R}_+$ 有

$$a^{\frac{1}{p}} b^{\frac{1}{q}} \leq \frac{a}{p} + \frac{b}{q}.$$

2. 设 $[a, b]$ 是给定的区间, $C[a, b]$ 是 $[a, b]$ 上全体连续函数的集, 则由

$$\rho(f, g) := \sup_{a \leq t \leq b} |f(t) - g(t)|, \qquad f, g \in C[a, b],$$

定义的 ρ 为 $C[a,b]$ 上的一个距离.

设 $\varphi \in C[a,b]$, 且对任何 $t \in [a,b]$, $\varphi(t) > 0$, $p \geqslant 1$, 则由

$$\rho_p(f,g) := \left[\int_a^b |f(t)-g(t)|^p \varphi(t)\, dt\right]^{\frac{1}{p}}, \quad f,g \in C[a,b],$$

定义的 ρ_p 也是 $C[a,b]$ 的距离.

3. 设 $\alpha := (\alpha_1, \alpha_2, \cdots)$ 为正实数序列, 满足 $\sum_{k=1}^{\infty} \alpha_k < \infty$. $p \geqslant 1$, $E := \mathbb{R}$ 或 $\mathbb{C}$ (全体复数集). 令

$$E^{\mathbb{N}}(\alpha;p) = \left\{x : x = (x_1, x_2, \cdots), x_k \in E, k \in \mathbb{N}, \sum_{k=1}^{\infty} \alpha_k |x_k|^p < \infty\right\},$$

$$\rho_p(x,y) = \left[\sum_{k=1}^{\infty} \alpha_k |x_k - y_k|^p\right]^{\frac{1}{p}}, \qquad x,y \in E^{\mathbb{N}}(\alpha;p).$$

则 $(E^{\mathbb{N}}(\alpha;p), \rho_p)$ 为一距离空间.

此外, 设

$$E^{\mathbb{N}} := \{x : x = (x_1, x_2, \cdots), \sup_{x \in \mathbb{N}} |x_k| < \infty\}$$

$$\rho(x,y) := \sup_{k \in \mathbb{N}} |x_k - y_k|, \ x,y \in E^{\mathbb{N}},$$

则 $(E^{\mathbb{N}}, \rho)$ 也是一距离空间. $\lim_{p\to\infty} \rho_p = \rho$ 是否还成立?

4. 设 $E := \mathbb{R}$ 或 $\mathbb{C}$.

$$E^{\mathbb{N}} := \{x : x = (x_1, x_2, \cdots), x_n \in E, n \in \mathbb{N}\},$$

$\{\alpha_n : n \in \mathbb{N}\}$ 为一可和的正数列, 试证: $(E^{\mathbb{N}}, \rho)$ 是距离空间, 其中 ρ 如下定义

$$\rho(x,y) := \sum_{n=1}^{\infty} \alpha_n \frac{|x_n - y_n|}{1 + |x_n - y_n|}, \quad x,y \in E^{\mathbb{N}}.$$

5. 设 $E := \{x : x = (x_1, x_2, \cdots), x_k \in \{0,1\}, k \in \mathbb{N}\}$, d 为 $\{0,1\}$ 上的离散距离, 试证: 如下定义的 ρ 是 E 上的距离,

$$\rho(x,y) := \sum_{k=1}^{\infty} \frac{1}{2^k} d(x_k, y_k), \quad x,y \in E.$$

6. 设 p 是一给定素数. 对每一给定的 $n \in \mathbb{N}$, 令 $U_p(n)$ 是能整除 n 的 p 的幂的最大指数 (即 $p^{U_p(n)}$ 能整除 n, 但 $p^{U_p(n)+1}$ 不能整除 n). 设 $x = \pm r/s$ 为有理数, $r, s \in \mathbb{N}$, 定义 $U_p(x) := U_p(r) - U_p(s)$. 试证:

1) $U_p(x), x \in \mathbb{Q}$, 是 $\mathbb{Q}$ 到 $\mathbb{Z}$ 的一个映射.

2) 对任何 $x, y \in \mathbb{Q}$, 令

$$d(x, y) := \begin{cases} 0, & x = y, \\ p^{-U_p(x-y)}, & x \neq y. \end{cases}$$

则 $d(x, y)$ 是 $\mathbb{Q}$ 上的一个距离 (此距离称为 p-adic 距离). 而

$$d(x, y) \leq \max(d(x, z), d(z, y)), \qquad \forall x, y, z \in \mathbb{Q}.$$

注 对给定的素数 p, 由 p-adic 距离定义的有理数基本列 (基本列的一般定义参看定义 3.1) 出发, 采用由绝对值定义距离的有理数基本列出发构造实数一样的方法, 也可以构造出一个完备数域 (对 p-adic 距离而言). 这个数域与 p 又有关, 称为 p-adic 域. p-adic 域是一种非阿基米得域, 近来发现它在理论物理中有用.

§2. 开集, 闭集

有了距离, 可以完全按数学分析的办法定义和处理极限, 连续等等, 但是为了尽量与一般拓扑理论的处理方法一致, 我们先讲开集与闭集.

1. 定义 设 (E, d) 为一距离空间, $a \in E, r \in \mathbb{R}_+ \setminus \{0\}$, 则分别称 $B(a, r) := \{x \in E : d(x, a) < r\}$, $\overline{B}(a, r) := \{x \in E : d(x, a) \leqslant r\}$, $S(a, r) := \{x \in E : d(x, a) = r\}$ 为以 a 为心, r 为半径的 **(开)球**, **闭球**, **球面**.

若 $G\subset E$, 且对任何 $x\in G$, 存在 $r=r(x)>0$, 使 $B(x,r)\subset G$, 则称 G 为 **开集**. 若 $F\subset E$, 而 $F^c=E\setminus F$ 为开集, 则称 F 为 **闭集**.

显然, 设 $a=(a_1,a_2,\cdots,a_d),b=(b_1,b_2,\cdots,b_d)\in\mathbb{R}^d$, 则在 1.3 的例 2 中定义的距离 d_2 (即欧氏距离) 之下, $\mathbb{R}^d$ 中的开区间 $(a,b):=\{(x_1,x_2,\cdots,x_d):a_k<x_k<b_k,k=1,2,\cdots,d\}$ 为开集, 闭区间 $[a,b]:=\{(x_1,x_2,\cdots,x_d):a_k\le x_k\le b_k,k=1,2,\cdots,d\}$ 为闭集.

易证: 任何球 $B(x,r)$ 是开集, 因为若 $y\in B(x,r)$, 则 $d(x,y)<r$, 而 $B(y,r-d(x,y))\subset B(x,r)$. 事实上, 对任何 $z\in B(y,r-d(x,y))$, $d(z,y)<r-d(x,y)$, 因而 $d(x,z)<r$, $z\in B(x,r)$.

任何闭球 $\overline{B}(x,r)$ 是闭集, 因为对任何 $y\notin\overline{B}(x,r)$, 则 $d(x,y)>r$, 于是 $B(y,d(x,y)-r)\subset(\overline{B}(x,r))^c$, 事实上, 若 $z\in B(y,d(x,y)-r)$, 则 $d(y,z)<d(x,y)-r$, 因而, 由引理 1.5 知

$$r<d(x,y)-d(y,z)=|d(x,y)-d(y,z)|\leqslant d(x,z),$$

即 $z\in(\overline{B}(x,r))^c$. 故 $(\overline{B}(x,r))^c$ 为开集, 因而 $\overline{B}(x,r)$ 为闭集.

关于开集, 有以下基本性质:

2. 定理 对给定的距离空间 (E,d),

1) $\varnothing$ 及 E 为开集.

2) 任意个开集之并是开集;

3) 有限个开集之交是开集.

注 用拓扑学的语言来说, 此定理即: 对距离空间 (E,d) 按定义 1 决定的开集类 τ 是空间 E 的一个拓扑, (E,τ) 是一拓扑空间. 实际上, 若空间 E 有一子集类 τ, 称为开集类, 它满足上面的 1)–3), 则称 (E,τ) 为具有拓扑 τ 的拓扑空间.

证明 1) 显然. 今设 $G_\alpha \subset E$, $\alpha \in I$ 为一开集类，令

$$G = \bigcup_{\alpha \in I} G_\alpha.$$

则对任何 $x \in G$, 必有一 $\alpha_0 \in I$, 使 $x \in G_{\alpha_0}$, 由 G_{α_0} 为开集知存在 $r > 0$, 使

$$B(x, r) \subset G_{\alpha_0} \subset G,$$

所以 G 为开集，故 2) 获证.

其次，设 $G_k \subset E$, $k = 1, 2, \cdots, n$ 为开集，则对任何 $x \in G := \bigcap_{k=1}^n G_k$, 必有 $x \in G_k$, $k = 1, 2, \cdots, n$, 因而对每一 $k = 1, 2, \cdots, n$, 有一 $r_k > 0$, 使 $B(x, r_k) \subset G_k$. 令 $r = \min_{1 \leq k \leqslant n} r_k$, 则 $r > 0$ 且

$$B(x, r) \subset B(x, r_k) \subset G_k, \quad k = 1, 2, \cdots, n$$

因而 $B(x, r) \subset \bigcap_{k=1}^n G_k = G$. 故 G 为开集，因而 3) 获证. □

由闭集是开集的余集及 de Morgan 法则立得

3. 定理 对给定的距离空间 (E, d),

1) $\varnothing$ 及 E 为闭集；

2) 任意个闭集的交为闭集 (因而球面为闭集，因为 $S(a, r) = \overline{B}(a, r) \cap (B(a, r))^c$).

3) 有限个闭集的并为闭集.

由定义立知

4. 定理 离散空间的每一子集都是开集且又是闭集.

距离空间 (E, d) 的子集不一定都是开集或闭集，为此引进

5. 定义 给定 (E, d), 对 $x \in A \subset E$, 若有一 $r > 0$ 使 $B(x, r) \subset A$, 则称 x 为 A 的 **内点** (所以开集 G 也可以定义为：**每一点都是其内点的集**); 设 $\varnothing \neq A \subset E$, 则称包含 A 的任一开集为 A 的

一个 **邻域**；当 $A=\{x\}$ 时，称 $\{x\}$ 的邻域为 x 的邻域（显然 x 是 A 的内点的充要条件是存在 x 的邻域 $N(x)\subset A$）；$A\subset E$，$A^\circ:=\{x\in A:\ x$ 是 A 的内点$\}$，称为 A 的 **内核**；$A^c=E\setminus A$ 的内点称为 A 的 **外点**；$(A^c)^\circ$ 称为 A 的 **外部**；$\overline{A}:=((A^c)^\circ)^c$ 称为 A 的 **闭包**；$\partial A:=\overline{A}\cap\overline{(A^c)}=\overline{A}\cap(A^\circ)^c=\overline{A}\setminus A^\circ$，称为 A 的 **边界**；∂A 的元称为 A 的 **边界点**.

这些概念从图形来看是很直观的，下面的一系列命题将给出它们的初步性质并确切说明它们的意义：

6. 定理 集合的内核有以下性质：

1) 对任何 $A\subset E$. 集合 A° 是 A 的最大开子集，即若 $B\subset A$ 为开集，则 $B\subset A^\circ$. 因而 A 是开集 $\iff A^\circ=A$；

2) $B\subset A\subset E\implies B^\circ\subset A^\circ$；

3) 对任何 $A,B\subset E$，则 $(A\cap B)^\circ=A^\circ\cap B^\circ$；

4) $A^\circ\cup B^\circ\subset(A\cup B)^\circ$，但等号未必成立.

证明留给读者.

还应注意：$A\neq\emptyset$，但可以有 $A^\circ=\emptyset$. 例如，A 为 $\mathbb{R}^2$ 中的一有限点集，则 $A^\circ=\emptyset$.

应用 de Morgan 法则即得下面的定理.

7. 定理 集的闭包有以下性质：

1) 对任何 $A\subset E$. $\overline{A}$ 是包含 A 的最小闭集，因而，A 是闭集 $\iff A=\overline{A}$；

2) $B\subset A\subset E\implies\overline{B}\subset\overline{A}$；

3) 对任何 $A,B\subset E\implies\overline{A\cup B}=\overline{A}\cup\overline{B}$；

4) 对任何 $A,B\subset E\implies\overline{A\cap B}\subset\overline{A}\cap\overline{B}$，但等号未必成立.

8. 定义 给定距离空间 (E,d), $A \subset E$, $x \in A$, 如果存在 x 的一个邻域 $N(x)$ 使得 $A \cap (N(x) \setminus \{4x\}) = \emptyset$, 则称 x 为 A 的 **孤立点**; $x \in E$, 如果对 x 的每个邻域 $N(x)$, $A \cap (N(x) \setminus \{x\}) \neq \emptyset$, 则称 x 为 A 的 **极限点** (limit point) 或 **聚点** (cluster point), A 的全体极限点组成的集称为 A 的 **导集**, 记作 A'.

显然, A 的极限点未必属于 A. 我们证明

9. 定理 给定距离空间 (E,d), $A \subset E$, 则 $A \cup A' = \overline{A}$, 因而 A 的任一极限点属于 A 的充要条件是 A 为闭集.

证明 若 $x \in \overline{A}$, 则 $x \notin (A^c)^\circ$, 因而对 x 的任何邻域 $N(x)$ 都有 $N(x) \bigcap A \neq \emptyset$. 若 $x \notin A$, 则 $(N(x) \setminus \{x\}) \bigcap A \neq \emptyset$, 因而 $x \in A'$, 若 $x \in A$, 显然有 $x \in A \cup A'$, 故 $x \in A \cup A'$. 反之, 设 $x \in A \cup A'$. 当 $x \in A$ 时, $x \in \overline{A}$; 当 $x \notin A$ 时, $x \in A'$, 因而对 x 的任一邻域 $N(x)$, 有 $(N(x) \setminus \{x\}) \bigcap A \neq \emptyset$, 故而 $x \notin (A^c)^\circ$, 即 $x \in \overline{A}$, 故无论哪种情况, 都有 $x \in \overline{A}$, 即 $A \cup A' \subset \overline{A}$. □

下列关于 $\mathbb{R}^d$ 中开集的结构性质, 在建立 $\mathbb{R}^d$上的有关测度时有用. 首先, 对于直线上的开集, 有下列很简洁的结果.

10. 定理 设 G 是直线上任一非空开集, 则存在开集族

$$\{(\alpha_k, \beta_k):\ -\infty \leqslant \alpha_k < \beta_k \leq \infty,\ k \in I\},$$

其中 $I = \{1, 2, \cdots, n\}$ 或 $\mathbb{N}$, (α_k, β_k), $k \in I$, 两两不交, 使

$$G = \bigcup_{k \in I} (\alpha_k, \beta_k).$$

证明的想法: 对任何 $x \in G$, 存在 (α, β), $\alpha < \beta$, $\alpha, \beta \in \mathbb{R}$ 使 $x \in (\alpha, \beta)$, 令

$$\alpha_x := \inf\{\alpha:\ x \in (\alpha, \beta) \subset G\}, \quad \beta_x := \sup\{\beta:\ x \in (\alpha, \beta) \in G\},$$

(α_x, β_x), $x \in G$ 中一切两两不同的开区间集即为所求. 详细证明请读者自己完成.

关于 $\mathbb{R}^d$, $d\in\mathbb{N}$, 有下列结果.

11. 定理 给定 $a_k:=(a_1^{(k)},\cdots,a_d^{(k)})\in\mathbb{R}^d$, $k=1,2$. 分别称

$$(a_1,a_2]:=\{(x_1,\cdots,x_d)\in\mathbb{R}^d:\ a_i^{(1)}<x_i\leqslant a_i^{(2)},i=1,2,\cdots,d\},$$

$$[a_1,a_2):=\{(x_1,\cdots,x_d)\in\mathbb{R}^d:\ a_i^{(1)}\leqslant x_i< a_i^{(2)},i=1,2,\cdots,d\}$$

为左开右闭, 左闭右开区间. $\mathbb{R}^d$中任一非空开集都能表成可列个两两不交的左开右闭 (左闭右开) 的正方体之并.

证明 1). 给定 $k\in\mathbb{N}$, 作以 $\dfrac{x}{2^k}$, $x\in 2\mathbb{Z}^d+e$ 为心 ($e:=(1,1,\cdots,1)$), 边长为 $1/2^{k-1}$ 且各边都与坐标轴平行的左开右闭正方体, 这种正方体称为第 k 级正方体, 它的全体组成的集记作

$$\mathcal{C}^{(k)}:=\{C_x^{(k)}:\ x\in 2\mathbb{Z}^d+e\},\quad k\in\mathbb{N}.$$

易见:

i) 对任意给定的 $k\in\mathbb{N}$,

$$\bigcup_{x\in 2\mathbb{Z}^d+e}C_x^{(k)}=\mathbb{R}^d;$$

ii) 给定 $k\in\mathbb{N}$, $C_x^{(k)}$, $x\in 2\mathbb{Z}^d+e$, 两两不交;

iii) 给定 k_1, $k_2\in\mathbb{N}$. $k_1>k_2$, $x_1,x_2\in 2\mathbb{Z}^d+e$, 则 $C_{x_1}^{(k_1)}\bigcap C_{x_2}^{(k_2)}=\varnothing$, $C_{x_1}^{(k_1)}\subset C_{x_2}^{(k_2)}$ 有且仅有一成立.

2). 首先将 $\mathcal{C}^{(1)}$ 中的元整个被 G 包含的正方体组成的集记作:

$$T^{(1)}:=\{C_x^{(1)}:\ x\in I_1\},\qquad I_1\subset 2\mathbb{Z}^d+e,$$

设 $T^{(1)},\cdots,T^{(k)}$ 取定, 则将 $\mathcal{C}^{(k+1)}$ 中的元整个被

$$G\setminus\left\{\Big(\bigcup_{x\in I_1}C_x^{(1)}\Big)\cup\cdots\cup\Big(\bigcup_{x\in I_k}C_x^{(k)}\Big)\right\}$$

包含的正方体组成的集记作:

$$T^{(k+1)}:=\{C_x^{(k+1)}:\ x\in I_{k+1}\},\qquad I_{k+1}\subset 2\mathbb{Z}^d+e\}.$$

于是由归纳法确定了 $T^{(k)}$, $k \in \mathbb{N}$. 今往证:

$$G = \bigcup_{k=1}^{\infty} \Big(\bigcup_{x \in I_k} C_x^{(k)} \Big).$$

3). 将上式中的右边的集合记作 G'. 显然 $G \supset G'$. 反之, 对任何 $x \in G$, 存在 $r > 0$, 使 $B(x,r) \subset G$. 由于 $\mathscr{C}^{(n)}$ 中的正方体的最长的对角线的长为 $\frac{\sqrt{d}}{2^{n-1}}$, 当 n 充分大时, 由 (1) 知必有一 $C_y^{(n)} \ni x$ 且由于二进制有理数的稠密性, $C_y^{(n)} \subset B(x,r) \subset G$. 而此 $C_y^{(n)}$ 必包含在 $\{C_x^{(k)}:\ x \in I_k,\ k \in \mathbb{N}\}$ 的某一正方形中, 因而 $x \in G'$. 故 $G \subset G'$. 从而定理证毕. □

习题

1. 给出 $\overline{B(x,r)} \neq \overline{B}(x,r)$ 的例子.

2. 给定 $a \in E$, 则任何 a 的邻域 A, 一定存在 $n \in \mathbb{N}$ 使 $B(a,\frac{1}{n}) \subset A$.

3. 设 $\varnothing \neq A \subset E$. 则 A 的任意有限个邻域的交仍然是 A 的邻域.

4. 证明: $A \subset E$ 为开集的充要条件是 $\partial A \cap A = \varnothing$.

5. 证明: $A \subset E$ 为闭集的充要条件是 $\partial A \subset A$.

6. 证明: $x \in \partial A$ 的充要条件是 x 的任何邻域既包含 A 的点又包含 A^c 的点.

7. 证明: $A \subset E$ 是开集的充要条件是对任意 $B \subset E$, 都有 $A \cap \overline{B} \subset \overline{A \cap B}$.

8. 给定距离空间 (E,d), $A \subset E$, 则

1) A' 为闭集;

2) $A \subset B \implies A' \subset B'$;

3) $(A \cap B)' \subset A' \cap B'$;

4) $(A \cup B)' = A' \cup B'$.

9. 作一实数列使其极限点集为空集.

10. 作一实数列使其极限点集为全体实数集.

11. 设 (E,d) 为距离空间. 给定 $A \subset E$. 试证: $\overline{A} \setminus A'$ 为 A 的全体孤立点组成的集.

12. 证明对任意 $x_0 \in \overline{A}$, 存在 $\{x_n\} \subset A$ 使 $d(x_n, x_0) \to 0$, $(n \to \infty)$.

§3. 完备性

在实数集中一个十分重要的性质是完备性 (completeness). 现在我们将它推广到距离空间上.

1. 定义 给定距离空间 (E,d). 称 $\{x_n : n \in \mathbb{N}\} \subset E$ **收敛** 于 $x \in E$, 记作 $x_n \to x$, $(n \to \infty)$. 如果对任何 $\varepsilon > 0$, 存在 $n_0 = n_0(\varepsilon)$, 使当 $n \geq n_0$ 时, 有

$$d(x_n, x) < \varepsilon.$$

$\{x_n : n \in \mathbb{N}\} \subset E$ 称为 E 中的 **基本列** (或 Cauchy 列). 如果对任何 $\varepsilon > 0$. 存在 $n_0 = n_0(\varepsilon)$, 有

$$d(x_n, x_m) < \varepsilon, \quad \forall\, n, m \geq 0. \tag{1}$$

显然, 若 $x_n \to x$. 则 $\{x_n\}$ 为基本列. 但是 (E,d) 中每一基本列是否收敛于 E 中某一点呢? 这便是 (E,d) 是否具有完备性的问题. 为此我们给出下列的

2. 定义 给定距离空间 (E,d). 如果对 E 中任何基本列 $\{x_n : n \in \mathbb{N}\}$ 都有一点 $x \in E$ 使 $x_n \to x$ $(n \to \infty)$. 则称 (E,d) 为 **完备空间** (complete space).

$A \subset E$, 则 d 在 A 上的限制 d_A 是 A 上的一个距离. 仍然记作 d, 并称 (A,d) 是 (E,d) 的子空间. 若 (E,d) 的子空间 (A,d) 完备, 则称 (A,d) 是 (E,d) 的 **完备子空间**.

本节所讨论的问题都是与距离有关的. 为了以后说话方便, 还引进一些概念和记号.

约定 $\sup\varnothing = 0$, $\inf\varnothing = \infty$. 给定距离空间 (E,d), $A\subset E$, 称

$$\delta(A) := \sup\{d(x,y):\ x,y\in A\}$$

为 A 的 **直径**, 若 $\delta(A)<\infty$, 则称 A 为 **有界集**. 若 $A, B\subset E$, 则称

$$d(A,B) := \inf\{d(x,y):\ x\in A,\ y\in B\}$$

为 **集合 A, B 间的距离**. 记 $d(x,B)=d(\{x\},B)$, 称为 **点 x 与集合 B 的距离**.

显然, A 为 (E,d) 中的有界集当且仅当存在 $a\in E$, $r\in(0,\infty)$, 使 $A\subset B(a,r)$. 而且还易证:

1) 若 $\varnothing\neq A\subset E$, 则对任何 $x,y\in E$, 有

$$|d(x,A)-d(y,A)|\leq d(x,y);$$

2) 对任何球 $B(a,r)$ 都有

$$\delta(B(a,r))\leq\delta(\overline{B}(a,r))\leq 2r;$$

3) 有界集的子集仍然有界;

4) (E,d) 的任何两个有界集 A, B 的并有界, 而且还有

$$\delta(A\cup B)\leq d(A,B)+\delta(A)+\delta(B).$$

证明 仅证 4) 中的不等式成立. 任取 $a\in A$, $b\in B$, 对一切 $x,y\in A\cup B$, 则或者它们都属于 A, 于是 $d(x,y)\leq\delta(A)$; 或者都属于 B, 于是 $d(x,y)\leq\delta(B)$, 或者 $x\in A$, $y\in B$, ($x\in B$, $y\in A$ 讨论方法相同), 则有 $d(x,y)\leq d(x,a)+d(a,b)+d(b,y)$. 故不论哪种情况, 都有

$$\delta(A\cup B)\leq\delta(A)+d(a,b)+\delta(B).$$

而此式对任何 $a\in A$, $b\in B$ 成立, 因而对 $d(a,b)$ 取下确界即得 4). □

3. 我们讨论在 §1 中所列出的各例的完备性.

例1 离散空间是完备的.

例2 $(\mathbb{R}^d, d_p)$ 完备，$(\mathbb{R}^d, d_\infty)$ 也完备，其中

$$d_\infty(x, y) := \max_k |x_k - y_k|.$$

例3 $(E^{\mathbb{N}}(p, \alpha), \rho_p)$ 完备.

证明 设 $x^{(n)} := (x_1^{(n)}, x_2^{(n)}, \cdots) \in E^{\mathbb{N}}(p, \alpha)$, $n \in \mathbb{N}$, 是基本列，则对任何 $k \in \mathbb{N}$,

$$\begin{aligned}|x_k^{(n)} - x_k^{(m)}| &\leq \frac{1}{\alpha_k^{\frac{1}{p}}}\left(\sum_{\ell=1}^{\infty} \alpha_\ell |x_\ell^{(n)} - x_\ell^{(m)}|^p\right)^{\frac{1}{p}} \\ &= \alpha_k^{\frac{1}{p}} \rho_p(x^{(n)}, x^{(m)}) \to 0, \quad (n, m \to \infty).\end{aligned}$$

因而由 $\mathbb{R}(\mathbb{C})$ 的完备性知存在 $x_k \in \mathbb{R}$, 使得

$$\lim_{n\to\infty} x_k^{(n)} = x_k.$$

今需证： $\sum_{n=1}^{\infty} \alpha_k |x_k|^p < \infty$, $\left(\sum_{k=1}^{\infty} \alpha_k |x_k^{(n)} - x_k|^p\right)^{\frac{1}{p}} \to 0$, $(n \to \infty)$.

由 $x^{(n)}$, $n \in \mathbb{N}$, 为基本列，从而对任何 $\varepsilon > 0$, 存在 N, 使当 $n, m > N$ 时，对任何 $M \in \mathbb{N}$ 都有

$$\sum_{k=1}^{M} \alpha_k |x_k^{(n)} - x_k^{(m)}|^p \leq \sum_{k=1}^{\infty} \alpha_k |x_k^{(n)} - x_k^{(m)}|^p < \varepsilon^p.$$

令 $m \to \infty$ 即得

$$\sum_{k=1}^{M} \alpha_k |x_k^{(n)} - x_k|^p \leq \varepsilon^p$$

由于 M 任意，令 $M \to \infty$, 即得对任何 $n \geq N$,

$$\left(\sum_{k=1}^{\infty} \alpha_k |x_k^{(n)} - x_k|^p\right)^{\frac{1}{p}} \leq \varepsilon,$$

再由 Minkowski 不等式得知对任何 $n \geq N$,

$$\left(\sum_{k=1}^{\infty} \alpha_k |x_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{\infty} \alpha_k |x_k^{(n)}|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} \alpha_k |x_k - x_k^{(n)}|^p\right)^{\frac{1}{p}}$$
$$\leq \rho_p(0, x^{(n)}) - \varepsilon < \infty$$

故 $x := (x_1, x_2, \cdots) \in E^{\mathbb{N}}(p, \alpha)$. 由 $\varepsilon > 0$ 任意即得 $\rho_p(x^{(n)}, x) \to 0$, 因而 $(E^{\mathbb{N}}(p, \alpha), \rho_p)$ 完备. □

例4 $(E^{\mathbb{N}}, \rho)$ 完备.

例5 $(C[a, b], \rho)$ 完备, 但 $(C[a, b], \rho_p)$ 不完备.

证明 前者证明显然, 后者主要因为极限可能不连续. 例如: 设

$$f_n(t) = \begin{cases} 0, & -1 \leq t < -\frac{1}{n}; \\ n(t + \frac{1}{n}), & -\frac{1}{n} \leq t < 0; \\ 1, & 0 \leq t \leq 1. \end{cases}$$

则

$$\rho_p(f_n, f_m) = \left[\int_{-1}^{1} |f_n(t) - f_m(t)|^p \, dt\right]^{\frac{1}{p}} \to 0.$$

令 $f(t) = \begin{cases} 1, & 0 \leq t \leq 1, \\ 0, & -1 \leq t < 0, \end{cases}$ 则 $\left[\int_{-1}^{1} |f_n(t) - f(t)|^p dt\right]^{\frac{1}{p}} \to 0$. 但 $f \notin C[-1, 1]$. 此外, 还可以证明对任何 $g \in C[-1, 1]$, 若

$$\rho_p(f_n, g) = \left[\int_{-1}^{1} |f_n(t) - g(t)|^p \, dt\right]^{\frac{1}{p}} \to 0,$$

则

$$\int_{-1}^{1} |f(t) - g(t)|^p \, dt = 0.$$

这是不可能的. □

注 这个例子启发我们：如果想要考虑对距离 ρ_p 完备的 $[a,b]$ 上的函数类，就必需考虑更大的函数类. 事实上，还必需推广积分的概念. 这个问题应用本书的第四，五两章的知识即可解决.

例6 $(\mathcal{D}_F, L)$ 完备.

暂不证，有兴趣者不妨试一试.

关于完备性有以下一些有用的简单性质 (证明留作习题).

4. 引理 在 (E,d) 中，基本列收敛当且仅当它有一收敛子列. $a \in \overline{A} \subset E$ 当且仅当存在序列 $\{a_n\} \subset A$, $a_n \to a$.

5. 引理 (E,d) 完备， (F,d) 是 (E,d) 的完备子空间当且仅当 F 是 (E,d) 的闭集.

数学分析中实数集 $\mathbb{R}$ 的完备性可以有很多种等价刻划，但现在在距离空间来说，这些刻划未必都等价. 相应于闭区间套引理，有下列的

6. 定理 若 (E,d) 完备，则任何直径趋于零的非空递降闭集列 (即 F_n 为闭集， $n \in \mathbb{N}$, 且 $F_n \supset F_{n+1}$, $n \in \mathbb{N}$) 的交不空；

反之，若 (E,d) 的任何直径趋于零的递降闭球列之交不空，则 (E,d) 完备.

证明 在每一 F_n 中取一点 x_n, 则对任何 $m \geq n$, $x_m \in F_n$, 且

$$d(x_n, x_m) \leq \delta(F_n) \to 0, \qquad (n \to \infty).$$

由 (E,d) 完备知，存在 $x \in E$, 使得 $d(x_n, x) \to 0 (n \to \infty)$. 由 $\{x_n, x_{n+1}, \cdots\} \subset F_n$, 应用引理 4 的后一结论知 $x \in \overline{F}_n = F_n$, 因而 $\bigcap_{n=1}^{\infty} F_n \supset \{x\} \neq \emptyset$.

反之，为证条件的充分性，用反证法，即由 (E,d) 不完备可构造出一半径趋于零的递降闭球列，而其交为空集.

由于 (E,d) 不完备，则存在 $\{x_n:\ n\in\mathbb{N}\}\subset E$ 为基本列，但没有极限. 现由此序列构造所需的闭球列. 设 n_1 是使

$$d(x_n,x_{n_1})<\frac{1}{2},\qquad \forall\, n\geq n_1$$

成立的足码. 设 $n_2>n_1$ 是使

$$d(x_n,x_{n_2})<\frac{1}{2^2},\qquad \forall\, n\geq n_2$$

成立的足码. 依次类推，对任何 $k\in\mathbb{N}$, 存在 $n_k\uparrow$ 使

$$d(x_n,x_{n_k})<\frac{1}{2^k},\qquad \forall\, n\geq n_k$$

成立. 则闭球列

$$\overline{B}(x_{n_1},1),\quad \overline{B}(x_{n_2},\frac{1}{2}),\cdots,\overline{B}(x_{n_k},\frac{1}{2^{k-1}}),\cdots$$

递降且半径趋于零. 事实上，对任何 $x\in\overline{B}(x_{n_{k+1}},\dfrac{1}{2^k})$, 有

$$d(x,x_{n_k})\leq d(x,x_{n_{k+1}})+d(x_{n_{k+1}},x_{n_k})\leq\frac{1}{2^{k+1}}+\frac{1}{2^k}<\frac{1}{2^{k-1}},$$

从而

$$x\in\overline{B}(x_k,\frac{1}{2^{k-1}}).$$

若 $\bigcap_{k=1}^{\infty}\overline{B}(x_{n_k},\dfrac{1}{2^{k-1}})\ni x$, 则当 $n\geq n_k$ 时,

$$d(x,x_n)\leq d(x,x_{n_k})+d(x_{n_k},x_n)<\frac{1}{2^{k-1}}+\frac{1}{2^k}\to 0,$$

即 $x_n\to x$, $(n\to\infty)$, 此与 $\{x_n\}$ 没有极限矛盾，所以

$$\bigcap_{k=1}^{\infty}\overline{B}(x_{n_k},\frac{1}{2^{k-1}})=\varnothing,$$

而这又与前提相矛盾，因而 (E,d) 完备. □

注 习知：实数集 $\mathbb{R}$ 是完备的，其中闭区间套引理成立. 由此可证：首项有界的每一递降的闭区间套的交集不空. 但是这个命题对一般的完备距离空间不成立. 我们在本节的习题 10 给出了一个反例，其证明作为读者的习题.

象从有理数构造实数一样, 对于不完备距离空间来说, 总可以用一种完全确定的方式把它包含在一个完备距离空间之中, 即将其完备化.

7. 定理 设 (E,d) 为任一距离空间, 则

1) 总存在完备距离空间 (E^*,d^*) 使 (E,d) 是 (E^*,d^*) 的子空间 (即 $E\subset E^*$ 且 $d^*(x,y)=d(x,y)$, 对任何 $x,y\in E$), 而且 $(\overline{E})_*=E^*$ (即 E 在 (E^*,d^*) 中的闭包为 E^*). ((E^*,d^*) 称为 (E,d) 的完备化)

2) (E,d) 的完备化是唯一的. 即若 (E^{**},d^{**}) 也是 (E,d) 的完备化, 则存在 E^* 到 E^{**} 的满射 φ 使:

(a) $$\varphi(x)=x,\qquad \forall\, x\in E;$$

(b) $\varphi(x^*)=x^{**},\ \varphi(y^*)=y^{**}\quad\Longrightarrow\quad d^{**}(x^{**},y^{**})=d^*(x^*,y^*)$

此时称 (E^*,d^*) 与 (E^{**},d^{**}) 等距).

定理的证明参考 [Su] 的有关部分.

习题

1. $\mathbb{R}$ 为实数集. 设

1).$\rho_1(x,y)=|\arctan x-\arctan y|$;

2).$\rho_2(x,y)=|e^x-e^y|$.

证明 $(\mathbb{R},\rho_1)$, $(\mathbb{R},\rho_2)$ 都不是完备距离空间.

2. 设 $\rho(m,n)=|m^{-1}-n^{-1}|$, $m,n\in\mathbb{N}$ 证明 $(\mathbb{N},\rho)$ 不完备.

3. $\mathbb{Z}$ 为整数集, $\rho(m,n)=|m-n|$, 证明 $(\mathbb{Z},\rho)$ 是完备距离空间.

4. 考虑三个定义在整个 $\mathbb{R}$ 上的连续函数集:

1) 有界连续函数集;

2) 满足条件 $\lim\limits_{x\to\infty}f(x)=0$ 的连续函数集;

3) 在有限区间外恒等于零的连续函数集.

如果规定 $\rho(f,g) = \sup\limits_{-\infty<x<-\infty} |f(x) - g(x)|$. 问哪一个是完备空间，哪一个不是？

5. 给定 $\Omega := \{\Delta : \Delta = [a,b], a < b, a,b \in \mathbb{R}\}$. 对于 $\Delta_1 = [a,b]$, $\Delta_2 = [c,d]$, 规定 $\rho(\Delta_1,\Delta_2) := |a-c| + |b-d|$, 证明： ρ 是 Ω 上的距离函数，但 (Ω,ρ) 不完备.

6. 若 (X_1,ρ_1) 与 (X_2,ρ_2) 等距同构， (X_1,ρ_1) 完备. 则 (X_2,ρ_2) 完备.

7. 证明 $(C[a,b],\rho)$ 完备.

8. 证明引理 4.

9. 证明引理 5.

10*. (Sierpinski 距离空间). 设 $E := \{x_n : n \in \mathbb{N}\}$, 试证：由

$$d(x_n,x_m) := \begin{cases} 1 + (n+m)^{-1}, & n \neq m, \\ 0, & n = m, \qquad n,m \in \mathbb{N}. \end{cases}$$

定义的 d 是 E 上的一个距离；(E,d) 的每一单点集都是开集，因而每一点都是孤立点； (E,d) 是完备距离空间.

再证： $\overline{B}(x_n, 1 + (2n)^{-1})$, $n \in \mathbb{N}$, 是一递降的闭球列，但它们的交集为空集. （此反例引自 [ST] 第 152 页，例 135.）

§4. 可分性，列紧性与紧性

可分性， 列紧性与紧性是实数集的三个重要性质， 也是距离
空间乃至拓扑空间的几个重要性质. 本节将就距离空间介绍这些概念和性质.

1. 定义 给定 (E,d). 若 $A\subset B\subset E$, $\overline{A}\supset B$, 则称 A **在** B **中稠** (dense). 若 $\overline{A}$ 不包含任何非空开集, 则称 A 为**疏朗集** (或**无处稠密集**). 若 E 有一可数稠子集, 则称 E **可分** (separable).

显然, 有理数集 $\mathbb{Q}$ 在 $\mathbb{R}$ 中稠且是可数集, 于是 $\mathbb{R}$ 可分. 由此, 还可知 $\mathbb{R}^n$ 可分, 事实上, $\mathbb{Q}^n := \{(x_1,\cdots,x_n): x_i\in\mathbb{Q}, i=1,2,\cdots,n\}$ 是它的一个可数稠子集.

还容易证明: 疏朗集的余集一定是稠集, 但其逆不真, 即稠集的余集不一定是疏朗集. 例如: 有理数集是稠集, 它的余集是无理数集, 也是稠集, 因而不是疏朗集. 以上的证明留作习题.

下面举两个著名的例子 —**Cantor 集和 Sierpinski 海绵**.

2. Cantor (康托) 集. 我们将介绍 Cantor 集的构造, 证明它是疏朗集, 完全集 (perfect set), (即 A 与它的导集 A' 相等), 并且它的势是 $\aleph$.

1) Cantor 集的构造. 将闭区间 $[0,1]$ 三等分, 去掉中间的开区间 $(\frac{1}{3},\frac{2}{3})$, 依同法将剩下的两个闭区间 $[0,\frac{1}{3}]$, $[\frac{2}{3},1]$ 各自三等分并去掉中间的开区间 $(\frac{1}{3^2},\frac{2}{3^2})$, $(\frac{2\cdot3+1}{3^2},\frac{2\cdot3+2}{3^2})$. 依此类推, 第 n 次去掉的 2^{n-1} 个开区间为:

$$\left(\frac{1}{3^n}\Big(\sum_{k=1}^{n-1}a_{n-k}3^k+1\Big),\frac{1}{3^n}\Big(\sum_{k=1}^{n-1}a_{n-k}3^k+2\Big)\right),$$

$$a_k=0 \text{ 或 } 2,\qquad k=1,2,\cdots,n-1.$$

将 $[0,1]$ 中的实数表示成三进制小数, 并令

$$\begin{aligned}G_0 :&= \bigcup_{n=1}^{\infty}\bigcup_{\substack{a_k=0,2\\k=1,\cdots,n-1}}\left(\frac{1}{3^n}\Big(\sum_{k=1}^{n-1}a_{n-k}3^k+1\Big),\frac{1}{3^n}\Big(\sum_{k=1}^{n-1}a_{n-k}3^k+2\Big)\right)\\&=\bigcup_{n=1}^{\infty}\bigcup_{\substack{a_k=0,2\\k=1,\cdots,n-1}}\big(0.a_1\cdots a_{n-1}1,\ 0.a_1\cdots a_{n-1}2\big).\end{aligned}$$

称 $P_0:=[0,1]\setminus G_0$ 为 **Cantor 集**.

2) 由于 G_0 是开集，所以 Cantor 集是闭集，而 $G_0 \cup (-\infty, 0) \cup (1, \infty)$ 是没有公共端点的两两不交的开区间的并，所以 P_0 没有孤立点 (即存在 $N(x)$ 使 $(N(x) \setminus \{x\}) \bigcap P_0 = \varnothing$). 因而由定理 2.9 ($\overline{A} = A \cup A'$) 知 $P_0 = (P_0)'$. 即 P_0 为完全集 (没有孤立点的闭集).

其次，我们可以将 $[0,1]$ 中的一切实数表为三进制的小数，而 G_0 的构成区间为

$$(0.a_1 \cdots a_{n-1}1,\ 0.a_1 \cdots a_{n-1}2),$$

$$a_k = 0 \text{ 或 } 2;\ k = 1, 2, \cdots, n-1;\ n \in \mathbb{N}$$

因此 G_0 中的实数表为三进制小数时，形如：

$$0.a_1a_2 \cdots a_{n-1}1, a_{n+1}a_{n+2} \cdots,$$

至少有一位数是 1, 而 $A := \{a \in [0,1] :\ a = 0.a_1a_2 \cdots a_k \cdots,\ a_k = 0 \text{ 或 } 2, k = 1, 2, \cdots\}$ 中必没有 G_0 的元，所以 $A \subset P_0 \subset [0,1]$. 设

$$B := \{(x_1, x_2, \cdots) :\ x_k = 0\text{或}1,\ k \in \mathbb{N}\},$$

由定理 1.4.3 知 $B \sim [0,1]$. 再令 $\varphi : A \mapsto B$ 为

$$\varphi(0.a_1a_2\cdots) = (b_1, b_2, \cdots),\ 0.a_1a_2\cdots \in A,$$

其中

$$b_k = \begin{cases} 0, & \text{若} a_k = 0, \\ 1, & \text{若} a_k = 2, \end{cases}$$

则显然有 $A \sim B$. 于是由对等关系的传递性得 $A \sim [0,1]$, 其势为 $\aleph$, 所以 P_0 的势也是 $\aleph$.

最后证明 P_0 是疏朗集. 用反证法，假设 P_0 包含某一开集，因而包含某一开区间 (α, β). 任取 $x \in (\alpha, \beta)$, 令 $\delta = \min\{x - \alpha, \beta - x\}$, 则 $\delta > 0$, 取 n_0 使得 $\dfrac{1}{3^{n_0}} < \delta$. 由于 $x \in P_0$ 是永远去不掉的点，因此 x 必属于第 n_0 次去掉开区间以后剩下的某一闭区间 $[a, b]$, 因而

$b-a=\frac{1}{3^{n_0}}$. 由 δ 和 n_0 的定义知 $[a,b]\subset(\alpha,\beta)$. 再对 $[a,b]$ 进行三等分时, 所去掉的中间的开区间记作 (α_0,β_0), 则 $(\alpha_0,\beta_0)\cap P_0=\varnothing$, 且 $(\alpha_0,\beta_0)\subset(\alpha,\beta)\subset P_0$. 这个矛盾就说明 P_0 不包含任何非空开集, 因而是疏朗集. □

3. Sierpinski 海绵 (Gasket). 从一个等边三角形 E_0 出发, 去掉以三边中点为顶点的三角形的内部, 留下来的集记作 E_1, 再对 E_1 中的每一个小三角形施行同样的手续, 剩下的集合记作 E_2. 如此继续下去, 得到 $E_3, E_4,\cdots$. 称集合 $Q_0:=\bigcap_{k=0}^{\infty}E_k$ 为 Sierpinski **海绵**. 它与 Cantor 集有着类似的性质. 有兴趣的读者可参考 [Fa1].

4. 定义 给定距离空间 (E,d). $A\subset E$ 的 **覆盖** 是指 E 的一组开子集, 它们的并包含 A. 若 $A\subset E$ 的任一覆盖都存在它的一个有限子类仍然是 A 的覆盖, 则称 A 为 **紧集**. 若 $A\subset E$ 的闭包 $\overline{A}$ 是紧集, 则称 A 为 **相对紧集**. 若 E 是紧集, 则称 E 为 **紧空间**. 若 E 的每一点都有一个相对紧的邻域, 则称 E 为 **局部紧空间**. 若 $A\subset E$ 的每一点列都有一收敛于 E 中点的子列, 则称 A 为 **列紧集**.

由数学分析知实数集 (在 1.3 的例 2 定义的距离之下) 的每一有限闭区间是紧集, 而实数集是局部紧空间但不是紧空间.

5. 定义 设 $A\subset E$. 如果对任何 $\varepsilon>0$, 存在 $n=n(\varepsilon)\in\mathbb{N}$ 及 $\{a_1,\cdots,a_n\}\subset A$, 使 $A\subset\bigcup_{k=1}^{n}B(a_k,\varepsilon)$. 则称 A 为 **全有界**.

6. 定理 下列三命题等价

1) E 可分;

2) E 有可数拓扑基 (即存在 E 的可数个开子集组成的集类 $\mathcal{G}$ 使得 E 的每一开子集都是 $\mathcal{G}$ 的某一子类的并, 此时称 $\mathcal{G}$ 为 E 的一个可数拓扑基);

3) E 的子集的任一覆盖都有可数子覆盖.

证明 1) $\Longrightarrow$ 2). 设 A 是在 E 中稠的可数子集，则 $\mathcal{G} := \{B(a, \frac{1}{m}) : a \in A, m \in \mathbb{N}\}$ 是 E 的一个可数拓扑基. 事实上，设 G 为 E 的开子集，令

$$G_1 := \bigcup_{B(a,\frac{1}{m})\subset G} B\Big(a, \frac{1}{m}\Big).$$

显然 $G_1 \subset G$, 反之 $x \in G$, 则存在 $r > 0$, 使 $B(x, r) \subset G$. 取 $m \in \mathbb{N}$, 使 $\frac{1}{m} \le \frac{r}{2}$, 则由 A 在 E 中稠可知：存在 $a \in B(x, \frac{1}{m})$, 因而 $x \in B(a, \frac{1}{m})$. 另一方面，对任何 $y \in B(a, \frac{1}{m})$,

$$d(x,y) \le d(x,a) + d(a,y) < \frac{1}{m} + \frac{1}{m} = \frac{2}{m} \le r,$$

故 $B(a, \frac{1}{m}) \subset B(x, r) \subset G$, 因而 $x \in G_1$, $G \subset G_1$.

2) $\Longrightarrow$ 3). 设 $\mathcal{G} := \{G_n : n \in \mathbb{N}\}$ 是 E 的可数拓扑基，$A \subset E$ 且 $\{V_\alpha : \alpha \in I\}$ 是 A 的一个覆盖，令

$$J := \{n \in \mathbb{N} : \exists\, \alpha \in I, \text{使 } G_n \subset V_\alpha\}.$$

对每一 $n \in J$, 取定 $\alpha_n \in I$, 使 $G_n \subset V_{\alpha_n}$, 则

$$\bigcup_{\alpha\in I} V_\alpha \supset \bigcup_{n\in J} V_{\alpha_n} \supset \bigcup_{n\in J} G_n.$$

由于 $\mathcal{G}$ 是拓扑基，若 $x \in V_\alpha$, 则存在 $n(x) \in J$ 使 $x \in G_{n(x)} \subset V_\alpha$. 于是 $x \in V_{\alpha_{n(x)}}$, $n(x) \in J$, 所以有

$$A \subset \bigcup_{\alpha\in I} V_\alpha = \bigcup_{n\in J} V_{\alpha_n}.$$

即 $\{V_{\alpha_n} : n \in J\}$ 是 A 的可数覆盖.

3) $\Longrightarrow$ 1) 对任何 $n \in \mathbb{N}$, $\{B(x, \frac{1}{n}) : x \in E\}$ 是 E 的覆盖，由 (3) 知 E 有可数覆盖 $\{B(x_{nk}, \frac{1}{n}) : k \in \mathbb{N}\}$, 于是，$\{x_{nk} : n, k \in \mathbb{N}\}$ 是 E 中可数稠子集，事实上，对任何 $x \in E$ 及 $\varepsilon > 0$, 存在 $n \in \mathbb{N}$ 满足 $\frac{1}{n} < \varepsilon$; 还存在 $k = k(n) \in \mathbb{N}$ 使 $x \in B(x_{nk}, \frac{1}{n})$, 因而 $d(x, x_{nk}) < \frac{1}{n} < \varepsilon$. $\square$

7. 引理 (E,d) 为紧空间的充要条件是：若它的每一闭集族的任意有限子族的交不空，则此闭集族的交不空.

由 de Morgan 法则即得此引理.

8. 引理 1) 紧空间的闭子集是紧集；

2) 距离空间的紧子集是闭集.

证明 设 $A\subset E$ 为闭集，E 为紧空间，$\{G_i: i\in I\}$ 是 A 的覆盖，则 A^c 是开集，且 $\{G_i: i\in I\}\bigcup\{A^c\}$ 是 E 的覆盖，因而存在 E 的有限覆盖，设为 $\{G_{i_1},\cdots,G_{i_n},A^c\}$. 于是 $\{G_{i_1},\cdots,G_{i_n}\}$ 为 A 的有限覆盖，故 1) 获证.

设 A 是紧集. 要证 2) 只需证：对任何 $x\notin A$, 存在 $r>0$, 使 $B(x,r)\cap A=\emptyset$(由此 $x\notin\overline{A}$, 因而 $\overline{A}\subset A$).

对任何 $y\in A$, 令 $r_y:=\dfrac{1}{2}d(x,y)>0$, 则 $B(x,r_y)\cap B(y,r_y)=\emptyset$. 而 $\{B(y,r_y): y\in A\}$ 是 A 的一个覆盖. 由 A 的紧性知存在 $\{y_1,\cdots,y_n\}\subset A$, 使 $A\subset\bigcup_{i=1}^n B(y_i,r_{y_i})$. 令 $r=\min_{1\le i\le n}r_{y_i}$, 则 $r>0$, 且 $B(x,r)\cap B(y_i,r_{y_i})=\emptyset$, $i=1,2,\cdots,n$, 从而有

$$B(x,r)\bigcap\Big(\bigcup_{i=1}^n B(y_i,r_{y_i})\Big)=\emptyset.$$

故有 $B(x,r)\cap A=\emptyset$, 2) 由此获证. □

注 此证明中关键是有分离性，所以此命题对 Hausdorff 空间 (即空间中任意两点存在不交的开邻域的拓扑空间) 成立.

9. 引理 距离空间的列紧集可分且全有界.

证明 设 A 为距离空间 (E,d) 的列紧集. 若 A 为有限集，结论自然成立，因此不妨设 A 为无限集.

今用归纳法证明. 首先任取 $x_1\in A$. 则 $d_1:=\sup_{y\in A}d(x_1,y)<\infty$, 否则存在 $\{y_n\}\subset A$, 使 $d(x_1,y_n)\to\infty$ $(n\to\infty)$, 由 A 列紧知存

在 $\{y_{n_k}\}$, 使 $y_{n_k} \to y$. $(k \to \infty)$. 因而 $d(x_1, y) = \lim_{k\to\infty} d(x_1, y_{n_k}) = \infty$, 这与 $d(x_1, y) < \infty$ 矛盾.

设 $\{x_1, \cdots, x_n\} \subset A$ 已取出，归纳地取 $x_{n+1} \in A$ 满足

$$\min_{1\leq k\leq n} d(x_k, x_{n+1}) \geq \frac{d_n}{2},$$
$$d_n := \sup_{y\in A} \min_{1\leq k\leq n} d(x_k, y).$$

则由归纳法可取出 $\{x_n: n \in \mathbb{N}\} \subset A$. 显然 $d_n \downarrow$. 往证 $d_n \to 0$ $(n \to \infty)$. 假设不然，则存在 $\varepsilon > 0$, 使得对任何 $n \in \mathbb{N}$, $d_n \geq \varepsilon$, 于是对任何 $n, m \in \mathbb{N}$ 有 $d(x_n, x_m) \geq \dfrac{\varepsilon}{2}$. 即 $\{x_n: n \in \mathbb{N}\}$ 的任何子列不是基本列，因而都不收敛，这与 A 列紧矛盾，所以 $d_n \to 0$. 因此对任何 $\varepsilon > 0$, 存在 $n_1 \in \mathbb{N}$, 使 $d_{n_1} < \varepsilon$. 故 $\sup_{y\in A} \min_{1\leq k\leq n_1} d(x_k, y) = d_{n_1} < \varepsilon$, 因而对任何 $x \in A$, 存在 $n \leqslant n_1$, 使 $d(x_n, x) < \varepsilon$, 即 $A \subset \bigcup_{n=1}^{n_1} B(x_n, \varepsilon)$, 这就说明 A 全有界. 同时也说明 $\{x_n: n \in \mathbb{N}\}$ 在 A 中稠，因而 A 可分. □

10. 定理 A 为紧距离空间 (E, d) 的紧子集当且仅当 A 为闭集且列紧.

证明 1) 设 A 是 E 的一个紧子集，则由引理 8.2) 知 A 是闭集.

假设 A 不列紧，则存在 $\{a_n\} \subset A$ 使它不包含收敛子列 (由于 A 是闭集，此事等价于不包含收敛于 A 中点的子列). 因此对任何 $x \in A$, 存在 $N(x)$ 只包含有 $\{a_n\}$ 的有限个点 (否则存在 $a \in A$, 使 a 的任一邻域包含 $\{a_n\}$ 的无限个点，因而 $\{a_n\}$ 有一子列收敛于 a), 而 $\{N(x): x \in A\}$ 是 A 的覆盖，由 A 的紧性知存在 $\{x_1, \cdots, x_k\} \subset A$ 使 $\bigcup_{i=1}^{k} N(x_i) \supset A \supset \{a_n\}$, 而 $\bigcup_{i=1}^{k} N(x_i)$ 只能包含 $\{a_n\}$ 的有限个点，所以 $\{a_n\}$ 只有有限个不同的点，因而它包含收敛的子序列，此矛盾即证明 A 为列紧的.

2) 设 A 为闭集且列紧，由引理 9 知 A 可分. 再由定理 6 3) 知 A 的任一覆盖都有可数子覆盖，所以只需证： A 的每一可数覆盖 $\{G_n: n\in\mathbb{N}\}$ 有一有限子覆盖. 反证法，设对任何 $n\in\mathbb{N}$, $\bigcup_{k=1}^n G_k \not\supset A$, 于是存在 $x_n\in A$ 使得 $x_n\notin\bigcup_{k=1}^n G_k$, 由 A 的列紧性及闭性知存在 $\{x_{n_k}\}$ 及 $x\in A$ 使 $x_{n_k}\to x$, $(k\to\infty)$. 另一方面，因为 $(\bigcup_{k=1}^n G_k)^c$ 为闭集，且随 n 的增加而下降，所以 $\{x_m: m\geq n\}\subset(\bigcup_{k=1}^n G_k)^c$. 故对任何 $n\in\mathbb{N}$, 有 $x\in(\bigcup_{k=1}^n G_k)^c$, 因而 $x\in(\bigcup_{k=1}^\infty G_k)^c\subset A^c$, 这个矛盾导致 $\{G_n\ n\in\mathbb{N}\}$ 中有 A 的一个有限子覆盖. □

11. 定理 设 K 是 $(E.d)$ 的一个子集，则下列三个命题等价：

1) K 列紧；

2) $\overline{K}$ 紧；

3) $\overline{K}$ 完备且 K 全有界.

而且每一紧距离空间完备可分全有界.

证明 1) $\Longrightarrow$ 2). 由定理 8 知只需证明： $\overline{K}$ 列紧. 设 $\{x_n\}\subset\overline{K}$, 则对任何 $n\in\mathbb{N}$, 存在 $y_n\in K$ 使 $d(x_n,y_n)<\frac{1}{n}$, 由 1) 知 $\{y_n\}$ 有一收敛子列 $\{y_{n_k}: k\in\mathbb{N}\}$ 收敛于 $y\in\overline{K}$, 显然 $\{x_{n_k}: k\in\mathbb{N}\}$ 也收敛于 y, 故 $\overline{K}$ 列紧.

2) $\Longrightarrow$ 3). 设 $\overline{K}$ 紧，则由 K 在 $\overline{K}$ 中稠知对任何 $\varepsilon>0$, $\{B(x,\varepsilon): x\in K\}$ 为 $\overline{K}$ 的覆盖，因而 $\overline{K}$ 有有限子覆盖 $\{B(x_k,\varepsilon): k=1,2,\cdots,n,\ x_k\in K\}$, 所以 K 全有界. 其次设 $\{x_n\}\subset\overline{K}$ 为基本列，又由定理 10 知 $\overline{K}$ 列紧，所以有 $\{x_n\}$ 的子列收敛于 $x_0\in\overline{K}$, 因而 $\{x_n\}$ 本身收敛于 $x_0\in\overline{K}$, 所以 $\overline{K}$ 完备.

3) $\Longrightarrow$ 1). 设 3) 成立，且 $\{x_n\}\subset K$. 若 $\{x_n\}$ 只有有限个不同点，则必有一收敛子列，因此不妨设 $\{x_n\}$ 两两不同. 由 K 全有界知对任何 $n\in\mathbb{N}$. 存在 $\{x_1^{(n)},\cdots,x_{k_n}^{(n)}\}\subset K$, 使 $\{B(x_\ell^{(n)},\frac{1}{n}): \ell=1,2,\cdots,k_n\}$ 为 K 的覆盖. 因此 $\{x_n\}$ 中有一无限子序列 $\{x_{1n}\}$

包含在某一 $B(x_{\ell_1}^{(1)},1)$ 中，此时 $(x_{\ell_1}^{(1)}\in K)$. 同理 $\{x_{1n}:\ n\in\mathbb{N}\}$ 有一无限子列 $\{x_{2n}\}$ 包含在某一 $B(x_{\ell_2}^{(2)},\frac{1}{2})$ 中，此时 $(x_{\ell_2}^{(2)}\in K)$. 依次类推，存在着可数个无限子列使

$$\{x_n\}\supset\{x_{1n}\}\supset\cdots\supset\{x_{kn}\}\supset\cdots,$$

而且 $\{x_{kn}\}$ 包含在一个半径为 $\frac{1}{k}$ 的球中. 取

$$\bar{x}_n:=x_{nn},\qquad n\in\mathbb{N}$$

则当 $n,m\geq N$ 时，$d(\bar{x}_n,\bar{x}_m)<\frac{2}{N}$ (因为都在 $B(x_{\ell_N}^{(N)},\frac{1}{N})$ 中). 即 $\{\bar{x}_n:\ n\in\mathbb{N}\}$ 为一基本列，由于 $\overline{K}$ 完备，$\{\bar{x}_n\}\subset\{x_n\}$ 收敛，因而 K 列紧，故 1), 2), 3) 等价.

为证最后的命题，设 (E,d) 为紧空间，则由 2), 3) 等价知 E 完备，由 1), 2) 等价 (或定理 10) 知 E 列紧且闭，再由引理 9 知 E 可分. □

注 在此定理的证明中，采用了下列的程序：先取序列 $\{x_n:\ n\in\mathbb{N}\}$ 的具有某种性质 (例如递减) 的可数个无限子列 $\{x_{kn}:\ n\in\mathbb{N}\}$, $k\in\mathbb{N}$, 然后令 $x_n:=x_{nn}, n\in\mathbb{N}$ 而得到适合某种要求的序列. 这种程序通常称为 **对角线程序** (diagonal procedure).

习题

1. 试证：$(\mathbb{R},d)$ $(d(x,y)=|x-y|)$ 完备，可分，但不是紧空间.

2. 试证：$([a,b],d)$ 为完备，可分，紧距离空间.

3. 试证：紧集的闭子集也是紧集，闭集是否一定是紧集？

4. 若 $\{A_n\subset E,\ n\in\mathbb{N}\}$ 是递降非空紧集序列，且 $d(A_n)=\sup_{x,y\in A_n}d(x;y)\to 0$, 则存在唯一的 $x_0\in\bigcap_{n=1}^{\infty}A_n$. 将 $\{A_n\}$ 改为闭集序列结论还成立吗？

5. 设 $A \subset B \subset E$, 若 $B \subset \overline{A}$, 则称 A 在 B 中稠, 试证: 设 $A, B, C \subset E$ 且 A 在 B 中稠, B 在 C 中稠, 则 A 在 C 中稠.

6. A 为疏朗集的充要条件是任何非空开集 B 都有一非空开集 $C \subset B$ 使 $C \cap A = \emptyset$.

7. 疏朗集的余集是稠集, 并举例说明稠集的余集不一定是疏朗集.

8. 给定 (E, d), $A \subset E$ 则

1) $a \in \overline{A}$ 当且仅当存在 $\{a_n: n \in \mathbb{N}\} \subset A$ 使 $d(a_n, a) \to 0$, (即 $a_n \to a$), $n \to \infty$.

2) $a \in A'$ 当且仅当存在 $\{a_n: n \in \mathbb{N}\} \subset A$, 且 a_n, $n \in \mathbb{N}$ 两两不同使 $d(a_n, a) \to 0$.

9. 设 (E, d) 为一紧距离空间, $\{G_\lambda: \lambda \in \Lambda\}$ 是 E 的一个覆盖, 则存在 $\alpha > 0$ 使任何半径为 α 的开球至少被包含在一个 G_λ 之中 (Lebesgue 性). 并试举一反例说明: 当 E 为全有界时, 上述结论不真.

10. 在紧距离空间中, 若 $\{x_n\}$ 两两不同, 且只有一个聚点 a, 则 $x_n \to a$ $(n \to \infty)$.

11. 相对紧集一定全有界, 而在完备距离空间中, 全有界集一定相对紧.

12. $\mathbb{R}^n$ 中的子集为紧集当且仅当它是有界闭集.

13. (E, d) 中任何两紧集之并仍为紧集.

14. 设 (E_n, d_n), $n \in \mathbb{N}$, 为一距离空间列.

$$E_\infty := E_1 \times E_2 \times \cdots := \{x: x = (x_1, x_2, \cdots), x_n \in E_n, n \in \mathbb{N}\}.$$

$d(x, x') := \sum\limits_{n=1}^{\infty} \frac{1}{2^n} \min(d_n(x_n, x_n'), 1)$, 试证:

1) (E_∞, d) 为一距离空间;

2) 设 $n \in \mathbb{N}$, U_k 是 E_k 的开子集 $(k = 1, 2, \cdots, n)$, 则 (E_∞, d) 中一切形如

$$G := U_1 \times \cdots \times U_n \times E_{n+1} \times \cdots$$
$$:= \left\{ x \in E_\infty : \begin{array}{l} x = (x_1, \cdots, x_n, \cdots), x_k \in U_k, k = 1, 2, \cdots, n; \\ x_\ell \in E_\ell, \ell = n + 1, n + 2, \cdots \end{array} \right\}$$

都是 E_∞ 的开集;

3) (E_∞, d) 的任一开集都是形如 (2) 的集的并, 即一切由 (2) 列出的集族是 (E_∞, d) 的一个拓扑基;

4) 若每一 (E_n, d_n) 可分, 则 (E_∞, d) 可分;

5) 若每一 (E_n, d_n) 是紧空间, 则 (E, d) 也是紧空间.

§5. 距离空间上的映射与函数

数学分析中的函数的连续性概念等都可以推广到距离空间, 甚至更广的空间. 为此先给出下列定义.

1. 定义 设 (E, d), (S, ρ) 为给定的距离空间, $f: E \mapsto S$ 为 E 到 S 的映射, 若对给定的 $x_0 \in E$ 及任何的 $\varepsilon > 0$, 存在 $\delta = \delta(x_0, \varepsilon) > 0$, 使当 $d(x, x_0) < \delta$ 时, 有 $\rho(f(x), f(x_0)) < \varepsilon$, 则称 f **在点** x_0 **连续**; 若 f 在每一 $x \in E$ 连续, 则称 f **在** E **上连续**; 若 f 为 E 到 S 的一一映射, 则逆映射 f^{-1} 存在; 若 f, f^{-1} 都连续, 则称 f 为 E 与 S 间的**同胚映射**; 若 f 为 E 到 S 的一一映射, 且 $\rho(f(x), f(y)) = d(x, y)$, $x, y \in E$, 则称 f 为 E 与 S 间的**等距映射**.

显然 f 等距、f 同胚、f 连续、f 在任何点 $x \in E$ 连续各概念依次一个蕴含后一个.

例 $f(x) := \dfrac{x}{1 + |x|}$, $x \in \mathbb{R}$ 是 $(\mathbb{R}, d)$ 与 $((-1, 1), \rho)$ 间的等距映射.

易知 f 是 $\mathbb{R}$ 到 $(-1,1)$ 的一一映射, 在 $\mathbb{R}$ 上定义 $d(x,y):=|x-y|$, 在 $(-1,1)$ 上定义 $\rho(x,y)=|f^{-1}(x)-f^{-1}(y)|$. 则对任何 $x,y\in\mathbb{R}$ 有 $\rho(f(x),f(y))=|f^{-1}(f(x))-f^{-1}(f(y))|=|x-y|=d(x,y)$. 因而 f 是 $(\mathbb{R},d)$ 到 $((-1,1),\rho)$ 上的等距映射.

若补充定义 $f(-\infty)=-1$, $f(\infty)=1$, 则将 f 扩充到 $\overline{\mathbb{R}}$ 上而成为 $\overline{\mathbb{R}}$ 到 $[-1,1]$ 的一一映射.

2. 引理 设 (E,d)、(S,ρ) 为距离空间且 $f:\ E\mapsto S$, 则下列各命题等价:

1) f 连续;

2) 对 S 的任何开子集 G, $f^{-1}(G)=\{x\in E:\ f(x)\in G\}$ 为 E 的开子集;

3) 对 S 的任何闭子集 F, $f^{-1}(F)$ 为 E 的闭子集;

4) 对一切 $A\subset S\implies f^{-1}(\overline{A})\supset\overline{f^{-1}(A)}$;

5) 对一切 $A\subset E\implies f(\overline{A})\subset\overline{f(A)}$.

证明 由 f 在 x_0 连续的定义知其等价于

$$(1)\qquad \forall\,\varepsilon>0,\ \exists\,\delta>0,\ 使\ f(B_d(x_0,\delta))\subset B_\rho(f(x_0),\varepsilon)$$

即 $B_d(x_0,\delta)\subset f^{-1}(B_\rho(f(x_0),\varepsilon))$.

若 (2) 成立, 则对一切 $x_0\in E$, $\varepsilon>0$, $f^{-1}(B_\rho(f(x_0),\varepsilon))$ 为开集, 由 x_0 属于此集知存在 $\delta>0$, 使 $B_d(x_0,\delta)\subset f^{-1}(B_\rho(f(x_0),\varepsilon))$. 因而 (1) 中的包含关系成立. 故 1) 成立.

反之, 设 f 连续且 G 是 S 的任一开集, 往证: $f^{-1}(G)$ 是开集. 对任何 $x\in f^{-1}(G)$ 有 $f(x)\in G$. 由于 G 是开集, 所以存在 $\varepsilon>0$ 使得 $B_\rho(f(x),\varepsilon)\subset G$. 再由 f 连续知 (即 (1)) 存在 $\delta>0$ 使得

$$f(B_d(x,\delta))\subset B_\rho(f(x),\varepsilon)\subset G,$$

故 $B_d(x,\delta)\subset f^{-1}(G)$, 因而 $f^{-1}(G)$ 是开集. 所以 1) 与 2) 等价.

2), 3), 4), 5) 的等价性的证明留作习题.

3. 引理 设 (E_1,d_1), (E_2,d_2), (E_3,d_3) 为距离空间, 且 f_1 是 E_1 到 E_2 的连续映射, f_2 是 E_2 到 E_3 的连续映射, 则 $f_2\circ f_1$ ($(f_2\circ f_1)(x):=f_2(f_1(x))$!) 为 E_1 到 E_3 的连续映射. (习题)

4. 引理 设 f,g 为距离空间 (E,d) 上的实 (复) 数值连续函数, ($f,g:\ E\mapsto\mathbb{R}(\mathbb{C})$), $\alpha\in\mathbb{R}$ ($\mathbb{C}$), 则

$$|f|:\ |f|(x):=|f(x)|;$$

$$\alpha f:\ (\alpha f)(x):=\alpha f(x);$$

$$f+g:\ (f+g)(x):=f(x)+g(x);$$

$$f\cdot g:\ (f\cdot g)(x):=f(x)\cdot g(x)$$

都是实 (复) 数值连续函数. 进而 (E,d) 上的全体实 (复) 数值连续函数对实 (复) 数乘法、函数加法及函数乘法作成一结合代数.

若 f,g 为数值连续函数, 且对一切 $x\in E$, $g(x)\neq 0$ 则

$$\frac{f}{g}:\ (\frac{f}{g})(x):=\frac{f(x)}{g(x)}$$

也是数值连续函数.

若 f,g 为实值连续函数, 则

$$f\vee g:\ (f\vee g)(x):=f(x)\vee g(x)=\max(f(x),g(x));$$

$$f\wedge g:\ (f\wedge g)(x):=f(x)\wedge g(x)=\min(f(x),g(x)).$$

是实值连续函数. (习题)

5. 引理 设 (E,d) 为距离空间.

1) 设 A,B 为 E 的两个闭子集, 且 $A\cap B=\emptyset$, 则存在 E 上的连续函数 f, 使 $f(A)=\{1\}$, $f(B)=\{0\}$ 且 $f(E)\subset[0,1]$.

2) 设 A, B 为 E 的两个闭子集, $A \cap B = \varnothing$, f, g 为 E 上连续实函数, 则存在 $h: E \to \mathbb{R}$ 连续且

$$h(x) = \begin{cases} f(x), & x \in A; \\ g(x), & x \in B; \end{cases} f \wedge g \leq h \leq f \vee g,$$

其中实数值函数 $h_1 \leq h_2$ 是指: 对一切 $x \in E, h_1(x) \leq h_2(x)$.

3) 设 A 为 E 的闭子集, $f: A \mapsto \mathbb{R}$, 有界连续函数, 则存在 $\tilde{f}: E \mapsto \mathbb{R}$ 连续, 使 $\tilde{f}(x) = f(x)$, 对一切 $x \in A$, 且 $\sup_{x \in E} |\tilde{f}(x)| = \sup_{x \in A} |f(x)|$.

证明 i) 因为 $d(\cdot, A)$, $d(\cdot, B)$ 都是 E 到 $\mathbb{R}$ 的连续函数 (因为 $|d(x, A) - d(y, A)| \leq d(x, y)$) 且

$$d(x, A) \begin{cases} = 0, & x \in A; \\ > 0, & x \notin A, \end{cases}$$

故由

$$f(x) := \frac{d(x, B)}{d(x, A) + d(x, B)} = \begin{cases} 0, & x \in B; \\ 1, & x \in A. \end{cases}$$

定义的 $f: E \to \mathbb{R}$ 满足要求 1).

ii) 由 $h(x) := \dfrac{d(x, A) g(x) + d(x, B) f(x)}{d(x, A) + d(x, B)}$ 定义的 $h: E \mapsto \mathbb{R}$ 满足要求 2).

iii) 若 $f = 0$, 则结论显然, 因此设 $f \neq 0$, 记 $f_0(x) := f(x), x \in A$, $m_0 = \sup_{x \in A} |f(x)|$, $A_0 := \{x \in A: f_0(x) \leq -m_0/3\}$, $B_0 := \{x \in A: f_0(x) \geq m_0/3\}$, 则 A_0, B_0 为 E 的闭集且 $A_0 \cap B_0 = \varnothing$, 由 2) 知存在 $F_0: E \mapsto \mathbb{R}$ 连续, 且 $F_0(A_0) = \{-m_0/3\}$, $F_0(B_0) = \{m_0/3\}$, $-m_0/3 \leq F_0 \leq m_0/3$, 令 $f_1(x) := f_0(x) - F_0(x)$, $x \in A$, 则 $f_1: A \mapsto \mathbb{R}$ 连续且 $m_1 := \sup_{x \in A} |f_1(x)| \leq \frac{2}{3} m_0$. 将上述对 f_0, m_0 的讨论过程用于 f_1, m_1, 而且此种讨论过程可以归纳地进行下去, 于是得到连续函数列 F_i, $i = 1, 2, \cdots$, 使得: 对任何

$i \in \mathbb{N}$, 有

$$-\frac{1}{3}\left(\frac{2}{3}\right)^i m_0 \le F_i(x) \le \frac{1}{3}\left(\frac{2}{3}\right)^i m_0, \qquad x \in E.$$

$$\left|f(x) - \sum_{i=0}^{n} F_i(x)\right| \le \left(\frac{2}{3}\right)^{n+1} m_0, \qquad x \in A.$$

这些性质使得 $\tilde{f}(x) := \sum_{i=0}^{\infty} F_i(x)$ 绝对一致收敛, 因而 $\tilde{f}: E \mapsto \mathbb{R}$ 连续且在 A 上与 f 相同 (与 $\mathbb{R}$ 上相应定理同样证明), 还有

$$|\tilde{f}(x)| \le \sum_{i=0}^{\infty} |F_i(x)| \le \sum_{i=0}^{\infty} \frac{1}{3}\left(\frac{2}{3}\right)^i m_0 = m_0.$$

因而 $\sup_{x\in E} |\tilde{f}(x)| = \sup_{x\in A} |f(x)| = m_0$. □

6. 定义 设 (E,d), (S,ρ) 都是距离空间, $A \subset E$, $f: A \mapsto S$. 若对任何 $\varepsilon > 0$, 存在 $\delta > 0$, 使当 $x, y \in A$ 且 $d(x,y) < \delta$ 时, 都有

$$\rho(f(x), f(y)) < \varepsilon.$$

则称 f 在 A 上一致连续. 若 $A = E$, 则称 f **一致连续**.

7. 引理 设 (E,d), (S,ρ) 为两个距离空间, 且 (S,ρ) 完备, $A \subset E$ 在 E 中稠, $f: A \mapsto S$ 一致连续, 则存在一个唯一的 $g: E \to S$ 一致连续, 并且 $g(x) = f(x)$, 对一切 $x \in A$.

证明 因为 $\bar{A} = E$, 所以对任何 $x \in E$, 存在 $\{a_n\} \subset A$, 使 $a_n \to x$. $\{a_n\}$ 为 E 中的基本列, 由 f 的一致连续性知 $\{f(a_n)\}$ 是 S 中的基本列, 所以存在 $g(x) \in S$, 使得: $\rho(f(a_n), g(x)) \to 0$, 即 $f(a_n) \to g(x)$. 可以证明 $g(x)$ 与 $\{a_n\}$ 的选取无关, 因而 $g(x)$, $x \in E$ 是 E 到 S 的函数. 且易证: g 一致连续并满足引理的要求. □

8. 引理 给定两个距离空间 (E,d), (S,ρ), 且 (E,d) 紧, 若 $f: E \mapsto S$ 连续, 则 f 一致连续.

证明 留给读者.

9. 引理 设 (E,d) 为紧距离空间, $f: E \to \mathbb{R}$ 连续, 则 f 达到最大值及最小值.

证明 留给读者.

习题

完成正文中引理 2, 3, 4, 8 和 9 的证明.

第三章 测度空间与概率空间

在数学分析积分 (特别是重积分) 理论中，面积 (或体积) 是基础概念. 它具有可加性：即设几何体 A、B 不相交且分别有体积 $V(A)$, $V(B)$, 则几何体 $A \cup B$ (将几何体看成组成它的点的集合) 的体积

$$V(A \cup B) = V(A) + V(B).$$

我们还知道概率也有可加性. 事实上, 还有很多客观事物具有可加性, 例如, 任何一种材料的重量 (质量); 导体所负载的电荷；物体所受的压力等等. 将这些可加性加以概括就成为测度的概念. 所谓测度就是在一个给定集合 Ω 的某一子集类 $\mathcal{C}$ 上定义的一个满足完全可加性的非负函数 $\mu(A)$、$A \in \mathcal{C}$, 即若 $A_n : n \in \mathbb{N}$ 是 $\mathcal{C}$ 中两两不交的元, 且 $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{C}$, 则

$$\mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

(更广一些的情况是 μ 的值可以是向量, 本书不加讨论). 为了研究 μ, 首先需要讨论 $\mathcal{C}$ 应具有哪些性质才适合做 μ 的定义域, 这是本章 §1 所讨论的内容；其次将介绍构作测度的方法以及概率和 $\mathbb{R}^n$ 上一些具体的测度 (勒贝格测度, 勒贝格 - 斯蒂尔捷斯测度等); 再次, 将讨论测度的一些基本性质；最后将介绍分数维及其测度.

§1. 集类

本节将介绍一些在测度和概率论中常遇到的一些集类，并讨

论它们的基本性质.

1. 定义 设 Ω 是一给定的非空集合. 它的一些子集组成的类 $\mathbb{S}$ 称为 Ω 的一个 **半集代数**, 如果它满足:

(i) $\Omega \in \mathbb{S}$, $\emptyset \in \mathbb{S}$;

(ii) 若 $A, B \in \mathbb{S}$, 则 $A \cap B \in \mathbb{S}$;

(iii) 若 $A, A_1 \in \mathbb{S}$, $A_1 \subset A$, 则存在 $\{A_2, \cdots, A_n\} \subset \mathbb{S}$, $A_1, \cdots, A_n$ 两两不交且使 $A = \bigcup_{k=1}^n A_k$.

显然, 由 (ii) 知半集代数对有限交封闭.

2. 引理 Ω 的子集类 $\mathbb{S}$ 是半集代数的充要条件是 (i),(ii) 及

(iii)' $A \in \mathbb{S} \implies$ 存在 $\{A_1, \cdots, A_n\} \in \mathbb{S}$ 它们两两不交, 且使 $A^c := \Omega \setminus A = \bigcup_{k=1}^n A_k$.

(注意: $A \in \mathbb{S}$ 时, A^c 未必属于 $\mathbb{S}$).

证明留给读者.

3. 例1 设 $\Omega := \mathbb{Z}_+$, $F_{k\ell} := \{n \in \mathbb{Z}_+ : k \le n < \ell\}$, $k \le \ell \in \mathbb{Z}_+$ 或 $\ell = \infty$, 则 $\mathbb{S} := \{F_{k\ell} : k \le \ell, k, \ell \in \mathbb{Z}_+$ 或 $\ell = \infty\}$ 为半集代数.

例2 设 $\Omega := \mathbb{R}^d$, $d \ge 1$. 则

$$\mathbb{S} := \left\{ (a, b] \subset \mathbb{R}^d : \begin{array}{l} a = (a_1, \cdots, a_d),\ b = (b_1, \cdots, b_d), \\ -\infty \le a_k \le b_k \le +\infty,\ k = 1, 2, \cdots, d \end{array} \right\}$$

是 $\mathbb{R}^d$ 的一个半集代数. (当 $b_k = +\infty$ 时, 理解 x_k 可取任意大的实数, 但不能取 $+\infty$, 对 $a_k = -\infty$ 亦同样理解). 即在实直线上, 一切左开右闭的区间 (有限或无限) 作成实直线的一半集代数, 在实平面上, 一切左开右闭的矩形 (有限或无限) 作成实平面的一个半集代数, 等等.

4. 定义 Ω 的子集类 $\mathcal{A}$ 称为它的 **集代数**, (亦称 **布尔代数**) 如果它满足:

(i) $\Omega \in \mathcal{A}$;

(ii) $A, B \in \mathcal{A} \implies A \cup B,\ A \cap B \in \mathcal{A}$;

(iii) $A \in \mathcal{A} \implies A^c := \Omega \setminus A \in \mathcal{A}$.

显然，Ω 的一切子集的类是 Ω 的一个集代数. 此外，$\{\varnothing, \Omega\}$、$\{\varnothing, A, A^c, \Omega\}$，$A \subset \Omega$，都是集代数.

5.引理 Ω 的子集类 $\mathcal{A}$ 是 Ω 的集代数的充要条件是下列诸组条件之一成立：

(I) 定义 4 中的 (i)，(iii) 及 (ii)' $A, B \in \mathcal{A} \implies A \cup B \in \mathcal{A}$;

(II) 定义 4 中的 (i)，(iii) 及 (ii)'' $A, B \in \mathcal{A} \implies A \cap B \in \mathcal{A}$;

(III) $\Omega \in \mathcal{A}$ 及 (iii)' $A, B \in \mathcal{A} \implies A \setminus B \in \mathcal{A}$.

证明 由 $A \setminus B = A \cap B^c$ 易知 (ii)，(iii) $\implies$ (iii)'. 故定义 4 $\implies$ (III). 反之，若 (iii)' 成立，则 $\Omega \in \mathcal{A}$ 且对任何 $A \in \mathcal{A}$，$A^c = \Omega \setminus A \in \mathcal{A}$，即 (iii) 成立. 若 $A, B \in \mathcal{A}$，则由 (iii)' $(A^c)^c = A$ 及

$$A \bigcup B = (A^c \bigcap B^c)^c, \qquad A \bigcap B = A \bigcap (B^c)^c = A \setminus B^c$$

知 (ii) 成立. 故 (III) $\implies$ 定义 4. 同时也证明了 (III) $\implies$ (II) 及 (I). 再由 $A \cup B = (A^c \cap B^c)^c$，$A \cap B = (A^c \cup B^c)^c$ 及 $(A^c)^c = A$ 知 (I) $\implies$ (II). 而由 $A \setminus B = A \cap B^c$ 及 (iii) 知 (II) $\implies$ (III). 故引理获证. □

6. 引理 设 $\mathcal{S}$ 是 Ω 的半集代数，则

$$\mathcal{A} := \{ \bigcup_{k=1}^{n} A_k : A_1, \cdots, A_n \in \mathcal{S} \text{ 且两两不交}, n \in \mathbb{N} \}$$
$$= \{ \bigcup_{k=1}^{n} A_k : A_1, \cdots, A_n \in \mathcal{S}, n \in \mathbb{N} \}$$

为**包含 $\mathcal{S}$ 的最小集代数** (即任一包含 $\mathcal{S}$ 的 Ω 的集代数必包含 $\mathcal{A}$). 称此集代数为**由 $\mathcal{S}$ 生成的集代数**，有时记作 $\mathcal{A}(\mathcal{S})$.

证明 显然, $\mathcal{A}\supset\mathcal{S}$. 再证: $\mathcal{A}$ 是 Ω 的集代数, 事实上, $\Omega\in\mathcal{S}\subset\mathcal{A}$. 因此, 由引理 5 的 (II) 知只需证: 若 $A,B\in\mathcal{A}$, 则 $A^c\in\mathcal{A}$ 且 $A\cap B\in\mathcal{A}$. 由 $\mathcal{A}$ 的定义知

$$A=\bigcup_{k=1}^{n}A_k,\qquad B=\bigcup_{k=1}^{m}B_k,$$

其中 $A_k\in\mathcal{S}$, $B_\ell\in\mathcal{S}$, $A_1,\cdots,A_n$ 两两不交, $B_1,\cdots,B_m$ 两两不交. 于是

$$A\bigcap B=\Big(\bigcup_{k=1}^{n}A_k\Big)\bigcap\Big(\bigcup_{\ell=1}^{m}B_\ell\Big)=\bigcup_{k=1}^{n}\bigcup_{\ell=1}^{m}(A_k\bigcap B_\ell).$$

由定义 1 知 $A_k\cap B_\ell\in\mathcal{S}$, 显然对 $k=1,\cdots,n$, $\ell=1,\cdots,m$ 两两不交. 故 $A\cap B\in\mathcal{A}$. 其次由引理 2 知对任何 $k\in\{1,2,\cdots,n\}$, 存在 $\{A_{k1},\cdots,A_{km_k}\}\subset\mathcal{S}$ 两两不交, 使 $A_k^c=\bigcup_{\ell=1}^{m_k}A_{k\ell}\in\mathcal{A}$, 因此

$$A^c=\bigcap_{k=1}^{n}A_k^c\in\mathcal{A},$$

故 $\mathcal{A}$ 为一集代数.

再证: $\mathcal{A}=\mathcal{A}(\mathcal{S})$. 由于 $\mathcal{A}(\mathcal{S})$ 是集代数, 它对有限并封闭, 因而 $\mathcal{S}$ 的有限个元的并属于 $\mathcal{A}$, 故 $\mathcal{A}(\mathcal{S})\supset\mathcal{A}$. 反之, 由于 $\mathcal{A}$ 为集代数, 且包含 $\mathcal{S}$, $\mathcal{A}(\mathcal{S})$ 为最小集代数, 所以 $\mathcal{A}\supset\mathcal{A}(\mathcal{S})$, 故 $\mathcal{A}=\mathcal{A}(\mathcal{S})$.

最后, 令 $\widetilde{\mathcal{A}}:=\{A\subset\Omega:\ A=\cup_{k=1}^{n}A_k,\ A_1,\cdots,A_n\in\mathcal{S}\}$, 显然 $\widetilde{\mathcal{A}}\supset\mathcal{A}$, 若 $A\in\widetilde{\mathcal{A}}$, 则存在 $\{A_1,\cdots,A_n\}\subset\mathcal{S}$, 使

$$A=\bigcup_{k=1}^{n}A_k.$$

由于 $\mathcal{A}$ 为集代数, 且 $\mathcal{A}\supset\mathcal{S}$, 因而 $\mathcal{A}\supset\{A_1,\cdots,A_n\}$, 故 $A\in\mathcal{A}$. 于是 $\widetilde{\mathcal{A}}\subset\mathcal{A}$, 因而 $\mathcal{A}=\widetilde{\mathcal{A}}$. □

7. 例1 设 $\Omega=\mathbb{R}^d$, $d\in\mathbb{N}$, $\mathcal{S}$ 由第 3 目中例 2 规定, 则

$$\mathcal{A}:=\{A\subset\mathbb{R}^d:\ A=\bigcup_{k=1}^{n}A_k,\ A_1,\cdots,A_n\in\mathcal{S},\ n\in\mathbb{N}.\}$$

为集代数.

在研究测度时, 不仅涉及到集运算对有限并, 有限交封闭的问题, 更重要的是涉及集运算对可数并, 可数交封闭的问题. 于是给出下列的定义.

8. 定义 称 Ω 的子集类 $\mathcal{F}$ 为 Ω 的 σ **代数** (σ–**域**), 如果它满足:

(i) $\Omega \in \mathcal{F}$;

(ii) 若 $A \in \mathcal{F}$, 则 $A^c \in \mathcal{F}$;

(iii) 若 $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, 则 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

显然, 任一 σ–代数必为一集代数, Ω 的一切子集构成的集类是一 σ–代数, 因此 Ω 的 σ–代数存在, 而且它包含 Ω 的任何 σ–代数, 此外 $\{\varnothing, \Omega\}$, $\{\varnothing, A, A^c, \Omega\}$, $A \subset \Omega$, 也都是 Ω 的 σ–代数. 由 de Morgan 法则立得下述

9.引理 $\mathcal{A}$ 是 Ω 的一个 σ 代数的充要条件是定义 8 中的 (i), (ii) 及

(iii)' 若 $A_n \in \mathcal{F}$, $n \in \mathbb{N}$. 则 $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$.

10. 引理 Ω 的任意一族 σ 代数的交仍为 Ω 的 σ 代数.

读者自证之.

11. 定理 设 $\mathcal{C}$ 是 Ω 的一个子集类, 则存在一个唯一的 Ω 的 σ 代数 $\sigma(\mathcal{C})$, 它包含 $\mathcal{C}$ 而且被包含 $\mathcal{C}$ 的任一 σ 代数所包含. $\sigma(\mathcal{C})$ 称为 **由** $\mathcal{C}$ **生成的** σ **代数**. 或 **包含** $\mathcal{C}$ **的最小** σ **代数**.

证明 由于 Ω 的一切子集构成的 σ 代数包含 $\mathcal{C}$, 所以包含 $\mathcal{C}$ 的 σ 代数类不空. 令 $\mathcal{A}_0$ 为包含 $\mathcal{C}$ 的一切 σ 代数的交, 由引理 10 知 $\mathcal{A}_0$ 是 Ω 的一个 σ 代数. 显然, 它具有定理结论所述的性质, 而且唯一.

12. 例1 设 $\Omega = \mathbb{R}^d$, $d \in \mathbb{N}$, $\mathcal{S}$ 为第 3 目例 2 中规定的集类，称 $\sigma(\mathcal{C})$ 为 d **维 Borel 域 (代数)**. 记作 $\mathcal{B}^d$ 或 $\mathcal{B}(\mathbb{R}^d)$ (当 $d = 1$ 时，常简记作 $\mathcal{B}$ 或 $\mathcal{B}(\mathbb{R})$). 显然

$$\mathcal{B}^d = \sigma(\mathcal{A}(\mathcal{S})).$$

此外由定理 2.2.11 知 $\mathcal{B}^d$ 包含 $\mathbb{R}^d$ 的一切开集的类 $\mathcal{O}^d$, 一切闭集的类，因而

$$\mathcal{B}^d = \sigma(\mathcal{O}^d).$$

例2 设 $\Omega = \mathfrak{C}$ 为全体复数集，令 $i = \sqrt{-1}$,

$$\mathcal{S} := \{ A : A = \{ z \in \mathfrak{C} : z = x + iy,\ (x, y) \in (a, b] \},\ a, b \in \mathbb{R}^2 \}.$$

则称 $\sigma(\mathcal{S})$ 为 **复** Borel **域**.

例3 设 (E, ρ) 为一距离空间，$\mathcal{O}$ 表示 E 的开集类，称 $\mathcal{B} := \sigma(\mathcal{O})$ 为 E **的 (关于** ρ **的)**Borel **域**. 记作 $\mathcal{B}(E, \rho)$ 或 $\mathcal{B}(E)$.

例4 称 $\overline{\mathbb{R}} := \mathbb{R} \cup \{ -\infty, \infty \}$ 为广义直线，其中 $\pm\infty$ 与 $\mathbb{R}$ 的运算规定如下：

$$x + (\pm\infty) = (\pm\infty) + x = \pm\infty, \quad \forall\, x \in \mathbb{R},$$

$$x \cdot (\pm\infty) = (\pm\infty) \cdot x = \begin{cases} \pm\infty, & \forall\, x > 0, \\ \mp\infty, & \forall\, x < 0, \end{cases}$$

$$(\pm\infty) + (\pm\infty) = \pm\infty,$$

$$(\pm\infty) \cdot (\pm\infty) = +\infty,$$

$$(\pm\infty) \cdot (\mp\infty) = -\infty,$$

$$0 \cdot (\pm\infty) = (\pm\infty) \cdot 0 = 0.$$

而和 $(+\infty)+(-\infty)$, $(-\infty)+(+\infty)$ 不加定义. 在 $\overline{\mathbb{R}}$ 上定义距离:

$$\rho(x,y):=\begin{cases}\left|\frac{x}{1+|x|}-\frac{y}{1+|y|}\right|, & x,y\in\mathbb{R},\\ \left|\frac{x}{1+|x|}-1\right|, & x\in\mathbb{R},y=+\infty,\\ \left|\frac{x}{1+|x|}+1\right|, & x\in\mathbb{R},y=-\infty,\\ 2 & x=+\infty,y=-\infty.\end{cases}$$

而在 $\overline{\mathbb{R}}^d:=\{\, x:\ x=(x_1,\cdots,x_d),\quad x_i\in\overline{\mathbb{R}},\quad i=1,2,\cdots,d\,\}$ 上定义距离:

$$\rho(x,y):=\Big(\sum_{i=1}^{d}\rho^2(x_i,y_i)\Big)^{1/2},$$

$$x=(x_1,\cdots,x_d),y=(y_1,\cdots,y_d)\in\overline{\mathbb{R}}^d.$$

若令

$$\mathcal{S}:=\{\,(a,b]:\ a,b\in\overline{\mathbb{R}}^d\,\}.$$

则

$$\sigma(\mathcal{S})=\mathcal{B}(\overline{\mathbb{R}}^d).$$

且有

$$\mathcal{B}(\overline{\mathbb{R}}^d)\cap\mathbb{R}^d=\mathcal{B}(\mathbb{R}^d).$$

此处及以后, 定义 Ω 的子集类 $\mathcal{C}$ 及 Ω 的子集 A 的交

$$\mathcal{C}\cap A:=\{C\cap A:\ C\in\mathcal{C}\}.$$

在有关 σ–代数的讨论中, 一个十分重要的工具是单调类定理. 我们在本节余下的部分加以介绍.

13. 定义 Ω 的子集类 Π 称为 π–系, 如果它对交封闭. Ω 的子集类 Λ 称为 λ–系 (或 Dynkin 类), 如果它满足

(i) $\Omega \in \Lambda$;

(ii) 对真差封闭：即 $A, B \in \Lambda$, $A \subset B \implies B \setminus A \in \Lambda$;

(iii) 对不降序列的并封闭：即 $\{ A_n : n \in \mathbb{N} \} \subset \Lambda$, $A_n \uparrow \implies \bigcup_{n=1}^{\infty} A_n \in \Lambda$.

显然半集代数是 π–系.

14. 引理 若 Ω 的子集类 $\mathcal{C}$ 同时为 λ–系及 π–系，则 $\mathcal{C}$ 为 σ–代数.

证明 由 $\mathcal{C}$ 为 π–系知它对交封闭，再由 $\mathcal{C}$ 是 λ–系知 $\Omega \in \mathcal{C}$, 且对任何 $A \in \mathcal{C}$, $A \subset \Omega$, 则由 (ii) 知 $A^c = \Omega \setminus A \in \mathcal{C}$. 最后若 $A_n \in \mathcal{C}$, $n \in \mathbb{N}$, 则 $B_n := \bigcup_{k=1}^{n} A_k$, $n \in \mathbb{N}$ 是不降序列，且由 $A_k^c \in \mathcal{C}$ 及 $\mathcal{C}$ 是 π–系知

$$B_n^c = \bigcap_{k=1}^{n} A_k^c \in \mathcal{C}.$$

因而 $B_n \in \mathcal{C}$. 由 (iii) 知 $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \in \mathcal{C}$, 故由定义 8 知 $\mathcal{C}$ 是 σ 代数. □

15. 定理 (集合形式的单调类定理) 设 Ω 的子集类 $\mathcal{C}$ 是 π–系，$\Lambda(\mathcal{C})$ 是包含 $\mathcal{C}$ 的最小 λ–系，则

$$\Lambda(\mathcal{C}) = \sigma(\mathcal{C}).$$

因而任何包含 $\mathcal{C}$ 的 λ–系 $\Lambda \supset \sigma(\mathcal{C})$.

注 包含任何子集类 $\mathcal{C}$ 的最小 λ–系 $\Lambda(\mathcal{C})$ 的存在且唯一性的证明与定理 11 的证明相同. 本节习题 16 是一个更一般的结论.

证明 首先注意 Ω 的任一 σ 代数 $\mathcal{F}$ 显然是 λ–系. 故 $\sigma(\mathcal{C})$ 是包含 $\mathcal{C}$ 的一个 λ–系，因而 $\sigma(\mathcal{C}) \supset \Lambda(\mathcal{C})$, 因之只需证 $\Lambda(\mathcal{C})$ 是 π–系即可，令

$$\mathcal{A}_A := \{ B \in \Lambda(\mathcal{C}) : A \cap B \in \Lambda(\mathcal{C}) \}.$$

显然，若能证明：对任何 $A \in \Lambda(\mathcal{C})$ 有 $\mathcal{A}_A = \Lambda(\mathcal{C})$ 即可，证明分三步：即证明：

1). 对任何 $A \in \Lambda(\mathcal{C})$, $\mathcal{A}_A$ 是一 λ–系；

2). 对任何 $A \in \mathcal{C}$, $\mathcal{A}_A = \Lambda(\mathcal{C})$;

3). 对任何 $A \in \Lambda(\mathcal{C})$, $\mathcal{A}_A = \Lambda(\mathcal{C})$.

1). 由 $\mathcal{A}_A$ 的定义即知 $\Omega \in \mathcal{A}_A$. 设 B、$C \in \mathcal{A}_A$, 且 $B \subset C$, 则 $A \cap C$, $A \cap B \in \Lambda(\mathcal{C})$ 且 $A \cap B \subset A \cap C$. 由 $\Lambda(\mathcal{C})$ 是 λ–系知 $A \cap (C \setminus B) = (A \cap C) \setminus (A \cap B) \in \Lambda(\mathcal{C})$, 因而由 $\mathcal{A}_A$ 的定义知 $C \setminus B \in \mathcal{A}_A$. 若 $\{A_n\} \subset \mathcal{A}_A$, $A_n \uparrow$, 则 $A \cap A_n \in \Lambda(\mathcal{C})$ 且 $A \cap A_n \uparrow$. 因而 $A \bigcap (\bigcup_{n=1}^{\infty} A_n) = \bigcup_{n=1}^{\infty} (A \bigcap A_n) \in \Lambda(\mathcal{C})$. 由 $\mathcal{A}_A$ 的定义即知 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}_A$, 故 $\mathcal{A}_A$ 是一 λ–系.

2). 设 $A \in \mathcal{C}$, 则 $\mathcal{A}_A \subset \Lambda(\mathcal{C})$, 于是要证 $\mathcal{A}_A \supset \Lambda(\mathcal{C})$, 由 (a) 知 $\mathcal{A}_A$ 是 λ–系知只要证： $\mathcal{A}_A \supset \mathcal{C}$ 即可. 设 $B \in \mathcal{C}$, 则由 $\mathcal{C}$ 是 π–系知 $A \cap B \in \mathcal{C} \subset \Lambda(\mathcal{C})$, 故 $\mathcal{C} \subset \mathcal{A}_A$.

3). 设 $A \in \Lambda(\mathcal{C})$, 则对任何 $B \in \mathcal{C}$, 由 2) 知 $A \in \mathcal{A}_B$, 由 $\mathcal{A}_B$ 的定义知 $A \cap B \in \Lambda(\mathcal{C})$, 因而再由 $\mathcal{A}_A$ 的定义知 $B \in \mathcal{A}_A$, 故 $\mathcal{C} \subset \mathcal{A}_A$, 由 1) 即知 $\Lambda(\mathcal{C}) \subset \mathcal{A}_A \subset \Lambda(\mathcal{C})$, 即 $\mathcal{A}_A = \Lambda(\mathcal{C})$.

故定理的前一部分获证. 至于第二个结论由前一结论立知.

此定理在测度论和概率论的理论研究中是一个重要的工具. 它的用法如下：通常是已知 $\mathcal{C}$ 中元具有性质 S, 需要证明 $\sigma(\mathcal{C})$ 中的元也具有性质 S, 为此令 $\Lambda := \{ B \subset \Omega : B\text{具有性质}S \}$ 于是 $\Lambda \supset \mathcal{C}$. 然后证明 $\mathcal{C}$ 是 π–系，再证 Λ 是 λ–系，由定理知 $\Lambda \supset \sigma(\mathcal{C})$, 即 $\sigma(\mathcal{C})$ 中的元都具有性质 S. 这种方法称为 λ–系方法.

习题

1. 试验证第 3 目例 2 中 $d = 1, 3$ 的情形.

2. 设 $\mathcal{S}$ 是 Ω 的半集代数，试证：

1) 若 $\{ A, A_1, \cdots, A_n \} \subset \mathcal{S}$, 且 $A_k \subset A$, $k = 1, 2, \cdots, n$ 两两不交, 则存在 $\{ A_{n+1}, \cdots, A_s \} \subset \mathcal{S}$, 使 $A_1, \cdots, A_s$ 两两不交, 且 $A = \bigcup_{k=1}^{s} A_k$.

2) 若 $\{ A_1, \cdots, A_n \} \subset \mathcal{S}$, 则在 $\mathcal{S}$ 中存在两两不交的有限个集 $B_1, \cdots, B_t$, 使每个 A_k 可以表成若干个 B_j 之并.

3. 称 Ω 的子集类 $\mathcal{S}$ 为半环, 如果它满足定义 1 中的 (ii)(iii). 试证:

1) $\emptyset \in \mathcal{S}$,

2) $\mathcal{S} := \{ (a, b] : a, b \in \mathbb{R}^d, a \le b \}$ 是 $\mathbb{R}^d$ 的半环.

4. 称 Ω 的子集类 $\mathcal{R}$ 为环 (或布尔环), 如果它满足 $A, B \in \mathcal{R} \Longrightarrow A \cup B, \quad A \setminus B \in \mathcal{R}$, 试证: Ω 的子集类 $\mathcal{R}$ 是环的充要条件是对并及真差封闭.

5. 如果将对称差看作集合间的加法运算 “+”, 将交看作是集合间的乘法运算 “·”, 则 1) Ω 的任一集代数 $\mathcal{A}$ 对 “+”,“·” 作成一个具单位元的可换环, 而且每个元都是幂等的 (即 $A \cdot A = A$); 2) Ω 的任一环 $\mathcal{R}$ 对 “+” 及 “·” 作成一幂等可换环.

6. 设 $\mathcal{S}$ 为 Ω 的一个半环, 则 $\{ A \subset \Omega : A = \bigcup_{k=1}^{n} A_k, \quad n \in \mathbb{N}, \{ A_1, \cdots, A_k \} \subset \mathcal{S}$两两不交 $\}$ 为包含 $\mathcal{S}$ 的最小环, 由此说明 $\{ A \subset \mathbb{R}^d : A = \bigcup_{k=1}^{n} (a_k, b_k], n \in \mathbb{N}, a_k, b_k \in \mathbb{R}^d, a_k \le b_k, k = 1, \cdots, n \}$ 是一个环 (由此可以看出环在表述上有方便之处).

7. 若 $A_k \subset \Omega, k = 1, 2, \cdots, n$, 两两不交, 且 $\bigcup_{k=1}^{n} A_k = \Omega$. 设 $\mathfrak{E} := \{ A_1, \cdots, A_n \}$. 试将 $\mathcal{A}(\mathfrak{E})$ 的全部元用 $\mathfrak{E}$ 的元表出, 请读者考查此 $\mathcal{A}(\mathfrak{E})$ 与 $\Omega_n = \{ 1, 2, \cdots, n \}$ 的一切子集作成的集代数之间的关系.

8. 设 $\mathfrak{E}$ 是 Ω 的任一子集类, 且 $\mathfrak{E}_1 := \{ A : A \in \mathfrak{E}$ 或 $A^c \in \mathfrak{E} \} \cup \{ \emptyset, \Omega \}$, $\mathfrak{E}_2 := \{ \bigcap_{k=1}^{n} A_k : A_k \in \mathfrak{E}_1, k = 1, 2, \cdots, n, n \in \mathbb{N} \}$, 则 $\mathcal{A} := \{ \bigcup_{i=1}^{m} B_i : B_i \in \mathfrak{E}_2, i = 1, 2, \cdots, m$, 两两不交, $m \in \mathbb{N} \}$

是包含 $\mathfrak{E}$ 的最小集代数 (即由 $\mathfrak{E}$ 生成的集代数 $\mathcal{A}(\mathfrak{E})$).

9. 设 Ω 为不可数集. 试证 (i) Ω 的一切有限集, 可数集以及它们的余集作成一个 σ 代数; (ii) Ω 的一切有限集, 可数集作成 Ω 的一个 σ 环 (即 Ω 的对可数并及差封闭的子集类).

10. 设 $\mathfrak{E}$ 是 Ω 的任意子集类, $A \in \sigma(\mathfrak{E})$. 则有 $\mathfrak{E}$ 的一个可列子类 $\mathfrak{D}$ 使 $A \in \sigma(\mathfrak{D})$.

11. 设 $\mathfrak{E} := \{ A_k, k = 1, 2, \cdots \}$, 其中 $A_k \subset \Omega$, $k = 1, 2, \cdots$, 两两不交. 试求 $\sigma(\mathfrak{E})$.

12. 设 A 是 Ω 的一个子集, $\mathfrak{E} := \{ B : A \subset B \subset \Omega \}$. 试指出 $\sigma(\mathfrak{E})$ 由哪些集组成.

13. 设 $\mathfrak{E}$ 是 Ω 的一个子集类, $\emptyset \neq A \subset \Omega$. 令

$$\mathfrak{E} \cap A := \{ B \cap A : B \in \mathfrak{E} \}$$

试证: $\mathcal{A}(\mathfrak{E}) \cap A$ 是 A 的子集类 $\mathfrak{E} \cap A$ 生成的 A 的集代数, $\sigma(\mathfrak{E}) \cap A$ 是 A 的子集类 $\mathfrak{E} \cap A$ 生成的 A 的 σ 代数.

14. 若 $\mathcal{A}$ 是集代数且对一切两两不交的集序列的并封闭, 则 $\mathcal{A}$ 是 σ 代数.

15. 设 $\mathfrak{E}$ 是 Ω 的一个子集类, 它含有 Ω 且对对称差与可列交两种运算封闭, 问它是不是一个 σ 代数?

16. 设 S 是一组集运算. 若 Ω 中的非空子集类 $\mathfrak{E}$ 对 S 中每一集运算都封闭, 则称 $\mathfrak{E}$ 为一 $S-$ 类. 试证:

1) 任意多个 $S-$ 类之交还是 $S-$ 类;

2) 设 $\mathfrak{E}$ 是 Ω 的一个子集类, 则存在一个唯一的包含 $\mathfrak{E}$ 的最小 $S-$ 类.

17. 称空间 Ω 中满足下述条件的集系 $\mathfrak{D}$ 为 $d-$ 系:

1) $\Omega \in \mathfrak{D}$;

2) $A, B \in \mathfrak{D}$且$A \cap B = \emptyset \implies B \cup A \in \mathfrak{D}$ (对不交并封闭);

3) $A \subset B$, $A, B \in \mathcal{D} \implies B \setminus A \in \mathcal{D}$ (对真差封闭).

试证： π–系上的最小 d– 系等于 π–系上的最小集代数. 从而包含某一 π–系的 d– 系必包含此 π–系生成的集代数.

18. Ω 的子集类 $\mathfrak{M}$ 称为 Ω 的单调类, 如果它满足: (1) 对不降集列的并封闭: 即 $A_n \in \mathfrak{M}$ 且 $A_n \uparrow, n \in \mathbb{N}, \implies \bigcup_{n=1}^{\infty} A_n \in \mathfrak{M}$. (2) 对不升序列的交封闭: 即 $A_n \in \mathfrak{M}$ 且 $A_n \downarrow, n \in \mathbb{N}, \implies \bigcap_{n=1}^{\infty} A_n \in \mathfrak{M}$. 试证:

1) 若 $\mathcal{A}$ 为 Ω 的集代数且为单调类, 则 $\mathcal{A}$ 为 σ 代数.

2) Ω 的任一子集 $\mathcal{C}$ 上的最小单调类是存在的.

3) 若 $\mathcal{A}$ 为 Ω 的集代数, 则包含 $\mathcal{A}$ 的最小单调类等于 $\sigma(\mathcal{A})$, 因而包含 $\mathcal{A}$ 的单调类必包含 $\sigma(\mathcal{A})$.

§2. 单调函数与测度的构作

本节将讨论由单调函数出发构作测度及测度扩张的问题. 为此先给出测度的概念, 再简单复习一下增函数的性质. 然后由右连续增函数出发构作出 $\mathbb{R}$ 中左开右闭区间组成的半集代数上的测度, 证明由一般半集代数 $\mathcal{S}$ 上的 σ–有限测度可唯一地扩张成 $\sigma(\mathcal{S})$ 上的 σ–有限测度.

1. 定义 设 $\mathcal{C}$ 是 Ω 的一个子集类, $\mu: \mathcal{C} \to \overline{\mathbb{R}}_+ := [0, +\infty]$ 且至少有一 $A \in \mathcal{C}$ 使 $\mu(A) < \infty$.

1) 若对任何 $A, B \in \mathcal{C}$, $A \cup B \in \mathcal{C}$, $A \cap B = \emptyset$ 都有

$$\mu(A \cup B) = \mu(A) + \mu(B),$$

则称 μ 为 $\mathcal{C}$ **上的可加测度**;

2) 若对任何 $n \in \mathbb{N}$, $A_k \in \mathcal{C}$, $k = 1, 2, \cdots, n$, 两两不交且 $\bigcup_{k=1}^{n} A_k \in \mathcal{C}$ 都有

$$\mu\Big(\bigcup_{k=1}^{n} A_k\Big) = \sum_{k=1}^{n} \mu(A_k),$$

则称 μ 为 $\mathcal{C}$ **上的有限可加测度**;

显然若 $\mathcal{C}$ 为集代数, 则 $\mathcal{C}$ 上可加测度一定是有限可加测度.

3) 若对任何 $A_n \in \mathcal{C}$, $n = 1, 2, \cdots$, 两两不交且 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{C}$ 都有

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n),$$

则称 μ 为 $\mathcal{C}$ **上的测度** (**或** σ–**可加测度**).

若对任何 $A \in \mathcal{C}$, $\mu(A) \in \mathbb{R}_+$, 则称上述相应各种测度为 **有限的**; 若对任何 $A \in \mathcal{C}$, 存在 $\{ A_n : n \in \mathbb{N} \} \subset \mathcal{C}$, 使 $\bigcup_{n=1}^{\infty} A_n = A$ 且 $\mu(A_n) < \infty$, $n \in \mathbb{N}$, 则称上述各相应测度为 σ–**有限的**.

若 $\mathcal{F}$ 为 Ω 的一个 σ 代数, 则称 $(\Omega, \mathcal{F})$ 为 **可测空间**, $A \in \mathcal{F}$ 称为 (Ω 中关于 $\mathcal{F}$ 的)heiti 可测集; 若 μ 为 $\mathcal{F}$ 上的测度, 则称 $(\Omega, \mathcal{F}, \mu)$ 为 **测度空间**; 若 μ 为 $\mathcal{F}$ 上的测度且 $\mu(\Omega) = 1$, 则称 μ 为 $\mathcal{F}$ **上的概率**(或 **概率测度**), 称 $(\Omega, \mathcal{F}, \mu)$ 为 **概率空间**(或 **概率场**).

显然, 当 $\mathcal{C}$ 为集代数或 σ 代数时, 若存在 $\{ A_n : n \in \mathbb{N} \} \subset \mathcal{C}$, 使 $\bigcup_{n=1}^{\infty} A_n = \Omega$, $\mu(A_n) < \infty$, $n \in \mathbb{N}$, 则 μ 即为 $\mathcal{C}$ 上的 σ–有限测度.

2. 设 $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R}) = \mathcal{B}^1$, μ 为 $\mathcal{B}^1$ 上的有限测度, 令 $F(x) := \mu((-\infty, x])$, $x \in \mathbb{R}$, 则 F 为 $\mathbb{R}$ 上右连续增函数, 且 $F(-\infty) = 0$, $F(+\infty) = \mu(\mathbb{R})$ (其中 $F(+\infty)$ 等记号的意义见第 3 目). 若 μ 是 $\mathcal{B}^1$ 上的 σ–有限测度且对任何有界集 A, $\mu(A) < +\infty$, 但不是有限测度, 则一般不能按上述方法定义增函数, 此时, 任取一 $a \in \mathbb{R}$ (不妨设 $a = 0$) 令

$$F(x) := \begin{cases} \mu((0, x]), & x \geq 0 \\ -\mu((x, 0]), & x < 0, \end{cases}$$

则 F 也是一个 $\mathbb{R}$ 上右连续增函数, 因此 $(\mathbb{R}, \mathcal{B}^1)$ 上在有界集上取

有限值的测度 μ 都可决定 $\mathbb{R}$ 上右连续的增函数，它们满足

$$F(b)-F(a)=\mu((a,b]),\qquad \forall\ -\infty\le a\le b\le +\infty.$$

而且由 μ 决定的右连续函数满足此性质者，最多差一个常数.

注 有的书中采用左连续增函数. 例如当 μ 为 $\mathcal{B}^1$ 上有限测度时，令 $\widetilde{F}(x)=\mu((-\infty,x))$.

为了使测度的构作与扩张的问题提法更直观一些，我们先简单复习一下不降函数 (增函数) 的性质，这些性质在很多方面是有用的而且有启发性，它们的证明在一般“概率论与数理统计”教科书中可以查到.

3. 称函数 $f:\mathbb{R}\to\mathbb{R}$ 为 **增函数**(或 **不降函数**)，如果对任何 $x_1,x_2\in\mathbb{R}$，有 $x_1<x_2 \implies f(x_1)\le f(x_2)$. 它具有以下性质：

(I) 对任何 $x\in\mathbb{R}$，$\lim_{t\uparrow x}f(t)=f(x-)$ 及 $\lim_{t\downarrow x}f(t)=f(x+)$ 存在且有限，而且 $f(x-)\le f(x)\le f(x+)$.

$f(-\infty):=\lim_{t\downarrow -\infty}f(t)$ 及 $f(+\infty):=\lim_{t\uparrow\infty}f(t)$ 存在，但前者可能是 $-\infty$，后者可能是 $+\infty$.

(II) 对任何 $x\in\mathbb{R}$，f 在 x 处连续当且仅当 $f(x-)=f(x+)$. 称 f (未必是增函数) 在 x 处有一 **跳跃**，如果 f 在 x 处的两个单侧极限 $f(x-)$，$f(x+)$ 存在而且不等，此时称 x 为 f 的一 **跳跃点**，$|f(x+)-f(x-)|$ 称为 f 在 x 处的 **跃度**，于是由 (I)，(II)(在 1.3.6 的例 3 中也曾讨论过) 即知

(III) 增函数的不连续点都是跳跃点，其不连续点最多有可数个.

增函数的跳跃点集可能有有限聚点，但这个聚点不一定是跳跃点，也就是说，跳跃点集未必是闭集，参见下例.

例1 设 $x_0 \in \mathbb{R}$. 定义

$$f(x)=\begin{cases}0, & \text{当 } x \le x_0-1\\ 1-\dfrac{1}{n}, & \text{当 } x_0-\dfrac{1}{n}\le x<x_0-\dfrac{1}{n+1}, \quad n\in\mathbb{N}\\ 1, & \text{当 } x\ge x_0\end{cases}$$

则点 x_0 是 f 的跳跃点集 $\{ x_0-1/n:\ n\geqslant 2 \}$ 的一个聚点，但 f 在 x_0 处连续.

例2 设给定互不相同的点列 $\{ a_n:\ n\in\mathbb{N} \}\subset\mathbb{R}$, $\{ b_n:\ n\in\mathbb{N} \}\subset\mathbb{R}_+\setminus\{0\}$ 且 $\sum_{n=1}^{\infty} b_n<\infty$. 定义

$$f(x):=\sum_{n=1}^{\infty} b_n\delta_{a_n}(x), \qquad x\in\mathbb{R},$$

其中

$$\delta_t(x):=\begin{cases}0, & x<t,\\ 1, & x\ge t.\end{cases}$$

易证： f 是增函数，其跳跃点集是 $\{ a_n:\ n\in\mathbb{N} \}$, 而 f 在点 a_n 的跃度为 b_n.

若 $\{ a_n:\ n\in\mathbb{N} \}$ 是有理数集 (或 $\mathbb{R}$ 的一个可数稠集), 上例说明增函数的跳跃点集可以在 $\mathbb{R}$ 上到处稠密，但是由性质 (III) 知增函数的跳跃点最多可数.

(IV) $\mathbb{R}$ 上的增函数除某些跳跃点的函数值外，由该函数在一稠集上的值确定. 确切地说，设 f_1 与 f_2 是 $\mathbb{R}$ 上的两个增函数, D 在 $\mathbb{R}$ 上稠，且

$$f_1(x)=f_2(x), \qquad \forall\, x\in D,$$

则 f_1, f_2 具有相同的跳跃点，在同一跳跃点上跃度相同，而且除了某些跳跃点外， f_1 与 f_2 的值都相等.

证明　对任何 $x\in\mathbb{R}$, 存在 $\{t_n\}\subset D$, $\{t'_n\}\subset D$, 使 $t_n\uparrow x$, $t'_n\downarrow x$(因为 D 在 $\mathbb{R}$ 中稠!). 于是

$$f_1(x-)=\lim_{n\to\infty}f_1(t_n)=\lim_{n\to\infty}f_2(t_n)=f_2(x-),$$
$$f_1(x+)=\lim_{n\to\infty}f_1(t'_n)=\lim_{n\to\infty}f_2(t'_n)=f_2(x+).$$

由此即得结论.

f_1, f_2 的不同仅在它们的 (共同) 跳跃点 x 上才可能出现, 而且 $f_1(x), f_2(x)$ 只能取区间 $[f_1(x-), f_1(x+)]=[f_2(x-), f_2(x+)]$ 中的值. 以后的讨论可以看到对于与测度以及其他有关问题相联系的增函数, 适当修改跳跃点处的值即可满足一定的要求. 例如对于测度来说, 将 f 修改成

(a) $$\tilde{f}(x):=f(x+),\qquad\forall\, x\in\mathbb{R},$$

或

(b) $$\tilde{f}(x):=f(x-),\qquad\forall\, x\in\mathbb{R}.$$

更加方便. 而对于 Fourier 分析来说, 将 f 修改成

(c) $$\tilde{f}(x):=\frac{1}{2}(f(x+)+f(x-)),\qquad\forall x\in\mathbb{R}.$$

较方便. 当然还可以有其他的修改.

不难证明下列命题

(V) 若 f 为 $\mathbb{R}$ 上的增函数, 则

$$\tilde{f}:\ \tilde{f}(x):=f(x+),\qquad\forall\, x\in\mathbb{R}.$$

是处处右连续的增函数, 而

$$\tilde{f}:\ \tilde{f}(x):=f(x-),\qquad\forall\, x\in\mathbb{R}$$

为处处左连续的增函数.

还可以证明更广的命题

(VI) 若 D 在 $\mathbb{R}$ 中稠， f 是 D 上的增函数，则

$$\tilde{f}:\ \tilde{f}(x) := \inf_{x<t\in D} f(t)$$

为处处右连续的增函数.

4. 命题 设 F 为 $\mathbb{R}$ 上的右连续增函数，则在半集代数

$$\mathcal{S} := \{\ (a,b]:\ -\infty \le a \le b \le +\infty\ \} \tag{1}$$

上有唯一测度 $\mu = \mu_F$ 存在使得

$$\mu((a,b]) = F(b) - F(a) \qquad a \le b,\ a,b \in \mathbb{R} \tag{2}$$

而且 μ 在有限区间上的值有限 (因而 σ–有限!).

证明 令 $F(\infty) = \lim_{x\to\infty} F(x)$, $F(-\infty) = \lim_{x\to-\infty} F(x)$,

$$\mu((a,b]) = F(b) - F(a), \qquad a,b \in \overline{\mathbb{R}},\ a \le b.$$

显然 μ 是定义在 $\mathcal{S}$ 上的非负集函数，而且易证 μ 是有限可加的. 事实上，若 $(a,b] = \bigcup_{k=1}^{n}(a_k,b_k]$, 而且 $(a_k,b_k]$, $k = 1,2,\cdots,n$, 两两不交，则不妨设右端各项皆非空集且

$$a = a_1 < b_1 = a_2 < b_2 = \cdots < b_{n-1} = a_n < b_n = b.$$

于是

$$\begin{aligned}\mu((a,b]) &= F(b) - F(a) = F(b_n) - F(a_1)\\ &= \sum_{k=1}^{n}(F(b_k) - F(a_k)) = \sum_{k=1}^{n}\mu((a_k,b_k]).\end{aligned}$$

下面用 $\mathbb{R}$ 中有限闭区间的紧性证明 μ 是 σ–可加的，设

$$(a,b] = \bigcup_{k=1}^{\infty}(a_k,b_k], \tag{3}$$

其中 $-\infty \le a < b \le +\infty$, $(a_k,b_k]$, $k \in \mathbb{N}$, 两两不交，往证

$$\mu((a,b]) = \sum_{k=1}^{\infty}\mu((a_k,b_k]). \tag{4}$$

首先注意当 $a = b$ 时，上式左右两边都是零，因而等式成立. 以下将就 $a < b$ 分几步证明之.

(i) 对任何 $n \in \mathbb{N}$, 由 (3) 知 $(a,b] \supset \bigcup_{k=1}^{n}(a_k,b_k]$, 由于 $(a_1,b_1]$, $\cdots$, $(a_n,b_n]$ 两两不交, 不妨设

$$a \le a_1 < b_1 \le a_2 < b_2 \le \cdots < b_{n-1} \le a_n < b_n \le b.$$

于是

$$\begin{aligned}\mu((a,b]) = F(b) - F(a) &\ge F(b_n) - F(a_1)\\ &= \sum_{k=1}^{n}[F(b_k) - F(a_k)] + \sum_{k=2}^{n}[F(a_k) - F(b_{k-1})]\\ &\ge \sum_{k=1}^{n}[F(b_k) - F(a_k)] = \sum_{k=1}^{n}\mu((a_k,b_k]).\end{aligned}$$

令 $n \to \infty$, 即得

$$\mu((a,b]) \ge \sum_{k=1}^{\infty}\mu((a_k,b_k]) \tag{5}$$

(ii) 为了证明与 (5) 反向的不等式, 先设 $-\infty < a < b < +\infty$, 任意取定 $\varepsilon > 0$, 使 $a+\varepsilon < b$, 由 F 的右连续性知, 对任何 $k \in \mathbb{N}$, 存在 $\delta_k > 0$, 使

$$F(b_k + \delta_k) - F(b_k) < \frac{\varepsilon}{2^k}, \tag{6}$$

于是, $\{ (a_k, b_k + \delta_k):\ k \in \mathbb{N} \}$ 是 $[a+\varepsilon, b]$ 的一个复盖. 由于 $[a+\varepsilon,b]$ 是 $\mathbb{R}$ 的紧子集, 所以存在 $N \in \mathbb{N}$ 使

$$[a+\varepsilon, b] \subset \bigcup_{k=1}^{N}(a_k, b_k+\delta_k). \tag{7}$$

由 (7) 知存在 $k_1 \le N$, 使 $a+\varepsilon \in (a_{k_1}, b_{k_1}+\delta_{k_1})$. 若 $b_{k_1}+\delta_{k_1} \in [a+\varepsilon, b]$, 则由 (7) 知存在 $k_2 \le N$, 使 $b_{k_1}+\delta_{k_1} \in (a_{k_2}, b_{k_2}+\delta_{k_2})$, 一般地, 若 $k_1, \cdots, k_\ell$ 已选定, 使得

$$b_{k_i}+\delta_{k_i} \in (a_{k_{i+1}}, b_{k_{i+1}}+\delta_{k_{i+1}}),\ i = 1, \cdots, \ell-1, b_{k_\ell}+\delta_{k_\ell} \le b, \tag{8}$$

则由 (7) 知存在 $k_{\ell+1} \le N$, 使

$$b_{k_\ell}+\delta_{k_\ell} \in (a_{k_{\ell+1}}, b_{k_{\ell+1}}+\delta_{k_{\ell+1}}).$$

如此选下去，由 (7) 利用归纳法知一定存在一 k_m 使

$$b \in (a_{k_m}, b_{k_m} + \delta_{k_m}). \tag{9}$$

故由 (9), (8), (6) 得

$$\begin{aligned}
F(b) - F(a+\varepsilon) &\le F(b_{k_m} + \delta_{k_m}) - F(a_{k_1}) \\
&= F(b_{k_m} + \delta_{k_m}) - F(b_{k_{m-1}} + \delta_{k_{m-1}}) + F(b_{k_{m-1}} + \delta_{k_{m-1}}) \\
&\quad - \cdots - F(b_{k_1} + \delta_{k_1}) + F(b_{k_1} + \delta_{k_1}) - F(a_{k_1}) \\
&= \sum_{\ell=1}^{m-1} \left[F(b_{k_{\ell+1}} + \delta_{k_{\ell+1}}) - F(b_{k_\ell} + \delta_{k_\ell})\right] \\
&\quad + \left[F(b_{k_1} + \delta_{k_1}) - F(a_{k_1})\right] \\
&\le \sum_{\ell=1}^{m} [F(b_{k_\ell} + \delta_{k_\ell}) - F(a_{k_\ell})] \\
&\le \sum_{k=1}^{\infty} \left[F(b_k + \delta_k) - F(b_k)\right] + \sum_{k=1}^{\infty} [F(b_k) - F(a_k)] \\
&\le \sum_{k=1}^{\infty} \mu((a_k, b_k]) + \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} \\
&= \sum_{k=1}^{\infty} \mu((a_k, b_k]) + \varepsilon.
\end{aligned} \tag{10}$$

在上式中令 $\varepsilon \to 0$, 则由 F 右连续知：当 $a, b \in \mathbb{R}$ 时，

$$\mu((a,b]) = F(b) - F(a) \le \sum_{k=1}^{\infty} \mu((a_k, b_k]). \tag{11}$$

(iii) 当 $a = -\infty$, $b < +\infty$ 时，对任何 $N \in \mathbb{N}$, 有

$$[-N, b] \subset \bigcup_{k=1}^{\infty} (a_k, b_k].$$

于是 (ii) 中直到 (10) 式的证明当 $a + \varepsilon$ 换成 $-N$ 后完全适用，所以有

$$F(b) - F(-N) \le \sum_{k=1}^{\infty} \mu((a_k, b_k]) + \varepsilon.$$

令 $\varepsilon \to 0$ 即得

$$F(b) - F(-N) \le \sum_{k=1}^{\infty} \mu((a_k, b_k]).$$

再令 $N \to \infty$, 即得 (11). 对于 $a \in \mathbb{R}$, $b = +\infty$ 以及 $a = -\infty$, $b = +\infty$ 可以同样处理. 故知 (11) 对任何 $-\infty \le a < b \le +\infty$ 成立, 与 (5) 式联合得知 μ 为 σ–可加, 即 μ 为 $\mathcal{S}$ 上的测度.

唯一性显然, 而且对有限区间 $(a, b]$, $F(b)$, $F(a)$ 有限, 所以 $\mu((a, b]) = F(b) - F(a)$ 有限. □

注 命题 4 可以推广到多维 $\mathbb{R}^d$ 的情形, 结论的形式和证明都是一样的, 但表达上要麻烦得多, 所以不再在此叙述. 读者如有兴趣, 可参看 [YWL] 第二章 §3.2 引理 2–4 (第 119–126 页) 与命题 4 相应的结论如下:

5. 命题 设 F 为定义在 $\mathbb{R}^n$ 上的有限实值函数, 且满足

1) 对任何 $a, b \in \mathbb{R}^n$, $a \le b$, 有

$$\triangle_{b,a} F := \sum_{k=0}^{n} (-1)^k \sum_{1 \le i_1 < \cdots < i_k \le n} F(b - \sum_{\ell=1}^{k} (b_{i_\ell} - a_{i_\ell}) e_{i_\ell}) \ge 0.$$

其中 e_i 为第 i 个坐标轴上的单位向量 (对上式符号不太习惯的读者可以注意 $b - \sum_{\ell=1}^{k} (b_{i_\ell} - a_{i_\ell}) e_{i_\ell}$ 为将 b 中的第 $i_1, \cdots, i_k$ 个坐标换成 a 的相应坐标而得到的点);

2) $F(x) = F(x_1, \cdots, x_n)$, $x \in \mathbb{R}^n$, 对每一自变量右连续,

则在 $\mathbb{R}^n$ 的半集代数

$$\mathcal{S}^n := \{ (a, b] : a \le b, \quad a, b \in \overline{\mathbb{R}}^n \}$$

上有唯一的测度 $\mu = \mu_F$ 存在, 使得:

$$\mu((a, b]) = \triangle_{b,a} F, \qquad a, b \in \mathbb{R}^n, \quad a \le b.$$

关于由增函数构作测度的问题, 因为有了命题 4, 5, 我们可以将问题提得更一般些, 即由半集代数 $\mathcal{S}$ 上的测度构作由它生成的 σ–代数 $\sigma(\mathcal{S})$ 上的测度的问题. 因此, 我们先给出:

6. 定义 设 $\mathcal{C}_1, \mathcal{C}_2$ 都是 Ω 的子集类，$\mathcal{C}_1 \subset \mathcal{C}_2$, μ_i 是 $\mathcal{C}_i$ 上的测度或有限可加测度，$i = 1, 2$. 如果对任何 $A \in \mathcal{C}_1$ 有

$$\mu_1(A) = \mu_2(A),$$

则称 μ_2 是 μ_1 **在 $\mathcal{C}_2$ 上的扩张**, 称 μ_1 是 μ_2 **在 $\mathcal{C}_1$ 上的限制**.

7. 定理 (测度扩张定理) 设 μ 为 Ω 的半集代数 $\mathcal{S}$ 上测度，则 μ 在 $\mathcal{S}$ 生成的 σ 代数 $\sigma(\mathcal{S})$ 上存在一个扩张.

若 μ 在 $\mathcal{S}$ 上 σ–有限，则 μ 在 $\sigma(\mathcal{S})$ 上的扩张唯一. 即若 μ_1, μ_2 是 $\mathcal{S}$ 在 $\sigma(\mathcal{S})$ 上的扩张，则对任何 $A \in \sigma(\mathcal{S})$ 都有：

$$\mu_1(A) = \mu_2(A).$$

测度扩张定理的证明在下面将分若干步来完成，这些步骤写成一些引理. 并且还将引入一些概念.

首先，我们将半集代数上的测度扩张到它生成的集代数上，进而得出半集代数上测度的一些性质.

8. 引理 设 μ 是半集代数 $\mathcal{S}$ 上的测度 (或有限可加测度)，则 μ 在 $\mathcal{A}(\mathcal{S})$ 上存在唯一的扩张 $\bar{\mu}$.

证明 由引理 1.6 知对任何 $A \in \mathcal{A}(\mathcal{S})$, 存在 $n \in \mathbb{N}$ 及 $B_k \in \mathcal{S}$, $k = 1, 2, \cdots, n$, 两两不交，使 $A = \bigcup_{k=1}^n B_k$, 令

$$\bar{\mu}(A) = \sum_{k=1}^n \mu(B_k),$$

则由此定义的 $\bar{\mu}$ 合理的，即 $\bar{\mu}(A)$ 与 A 的表示法无关. 事实上，若 $A = \bigcup_{i=1}^m C_i$, $C_i \in \mathcal{S}$, $i = 1, \cdots, m$, 两两不交，则由 μ 在 $\mathcal{S}$ 上有限可加，得：

$$\sum_{k=1}^n \mu(B_k) = \sum_{k=1}^n \mu(B_k \cap A) = \sum_{k=1}^n \sum_{i=1}^m \mu(B_k \cap C_i)$$
$$= \sum_{i=1}^m \sum_{k=1}^n \mu(B_k \cap C_i) = \sum_{i=1}^m \mu(A \cap C_i) = \sum_{i=1}^m \mu(C_i).$$

$\tilde{\mu}$ 的非负性显然. 今往证: 若 μ 是测度, 则 $\tilde{\mu}$ 具有 σ–可加性. 设 $A=\bigcup_{n=1}^{\infty} B_n \in \mathcal{A}(\mathbb{S})$, $B_n \in \mathbb{S}$, $n\in\mathbb{N}$, 两两不交, 则由引理 1.6 知

$$A=\bigcup_{i=1}^{L} A_i,\quad A_i\in\mathbb{S},\quad i=1,2,\cdots,L,\quad \text{两两不交};$$

$$B_n=\bigcup_{j=1}^{J_n} B_{nj},\quad B_{nj}\in\mathbb{S},\quad j=1,2,\cdots,J_n, \text{两两不交},\quad n\in\mathbb{N}.$$

于是

$$C_{nij}:=A_i\cap B_{nj},\quad i=1,2,\cdots,L,\quad j=1,2,\cdots,J_n,\quad n\in\mathbb{N},$$

为 $\mathbb{S}$ 中两两不交的元, 故由 μ 的 σ–可加性及 $\tilde{\mu}$ 的定义知

$$\begin{aligned}\tilde{\mu}(A)&=\sum_{i=1}^{L}\mu(A_i)=\sum_{i=1}^{L}\mu\Big(\bigcup_{n=1}^{\infty}\bigcup_{j=1}^{J_n}C_{nij}\Big)=\sum_{i=1}^{L}\sum_{n=1}^{\infty}\sum_{j=1}^{J_n}\mu(C_{nij})\\&=\sum_{n=1}^{\infty}\sum_{j=1}^{J_n}\mu\Big(\bigcup_{i=1}^{L}C_{nij}\Big)=\sum_{n=1}^{\infty}\sum_{j=1}^{J_n}\mu(B_{nj})=\sum_{n=1}^{\infty}\tilde{\mu}(B_n).\end{aligned}$$

故 $\tilde{\mu}$ 在 $\mathcal{A}(\mathbb{S})$ 上 σ–可加, 用同法可由 μ 的有限可加导出 $\tilde{\mu}$ 有限可加, 而且表达上更简单些. □

9. 引理 设 μ 是半集代数上的有限可加测度, 若 $\{A,A_1,\cdots,A_n\}\subset\mathbb{S}$, $\bigcup_{k=1}^{n}A_k\subset A$ 且 A_k, $k=1,\cdots,n$ 两两不交, 则:

$$\sum_{k=1}^{n}\mu(A_k)\le\mu(A).$$

证明 考查 μ 在 $\mathcal{A}(\mathbb{S})$ 上的扩张 $\tilde{\mu}$, 令

$$A_{n+1}:=A\setminus\Big(\bigcup_{k=1}^{n}A_k\Big),$$

则 $A_1,\cdots,A_{n+1}$ 为 $\mathcal{A}(\mathbb{S})$ 中两两不交的元, 于是由 $\tilde{\mu}$ 是 μ 的扩张知:

$$\mu(A)=\sum_{i=1}^{n+1}\tilde{\mu}(A_i)=\sum_{i=1}^{n}\mu(A_i)+\tilde{\mu}(A_{n+1})\ge\sum_{i=1}^{n}\mu(A_i).\quad\square$$

10. 引理 1) 若 μ 是半集代数 $\mathcal{S}$ 上的有限可加测度，$\{A, A_1, \cdots, A_n\} \subset \mathcal{S}$ 且 $A \subset \bigcup_{k=1}^{n} A_k$，则

$$\mu(A) \leq \sum_{k=1}^{n} \mu(A_k). \qquad \text{(次可加性)}$$

2) 若 μ 为半集代数 $\mathcal{S}$ 上的测度，$A, A_n \in \mathcal{S}$，$n \in \mathbb{N}$，且 $A \subset \bigcup_{n=1}^{\infty} A_n$，则

$$\mu(A) \leq \sum_{n=1}^{\infty} \mu(A_n) \qquad \text{(次 } \sigma\text{–可加性)}$$

证明 1)、2) 的证法类似，只证 2)．1) 留给读者自己证明．为此考虑 μ 在 $\mathcal{A}(\mathcal{S})$ 上的扩张 $\tilde{\mu}$．首先证明 $A = \bigcup_{n=1}^{\infty} A_n$ 的情形．此时：

$$B_n := A_n \setminus \left(\bigcup_{k=1}^{n-1} A_k\right) \in \mathcal{A}(\mathcal{S}), \qquad n \in \mathbb{N}$$

两两不交，且

$$A = \cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} B_n$$

于是由 $\tilde{\mu}$ 是 μ 的扩张 (引理 8) 知：

$$\mu(A) = \sum_{n=1}^{\infty} \tilde{\mu}(B_n) \leq \sum_{n=1}^{\infty} \tilde{\mu}(A_n) = \sum_{n=1}^{\infty} \mu(A_n).$$

其中用到以下事实：$\tilde{\mu}(B_n) \leq \tilde{\mu}(A_n)$，$n \in \mathbb{N}$．实际上，由 $B_n \subset A_n$ 知

$$\begin{aligned}\tilde{\mu}(A_n) &= \tilde{\mu}(B_n \cup (A_n \setminus B_n)) \\ &= \tilde{\mu}(B_n) + \tilde{\mu}(A_n \setminus B_n) \geq \tilde{\mu}(B_n).\end{aligned}$$

若 $A \subset \bigcup_{n=1}^{\infty} A_n$，则 $A = \bigcup_{n=1}^{\infty}(A \cap A_n)$．且 $A \cap A_n \in \mathcal{S}$，$n \in \mathbb{N}$，于是由前面证明知

$$\mu(A) \leq \sum_{n=1}^{\infty} \mu(A \bigcap A_n) \leq \sum_{n=1}^{\infty} \mu(A_n). \quad \square$$

11. 定义 设 μ^* 为定义在 Ω 的一切子集类上的非负广义实函数且具有

1) $\mu^*(\emptyset)=0$;

2) 不降性：对任何 $A\subset B\subset\Omega$, 有 $\mu^*(A)\le\mu^*(B)$;

3) 次 σ–可加性：对任何 $A_n\subset\Omega$, $n\in\mathbb{N}$, 有

$$\mu^*\Big(\bigcup_{n=1}^{\infty}A_n\Big)\le\sum_{n=1}^{\infty}\mu^*(A_n)$$

则称 μ^* 为 Ω 的一个 **外测度**.

12. 引理 设 μ 为 Ω 的半集代数 $\mathcal{S}$ 上的测度，对任何 $A\subset\Omega$, 令

$$\mu^*(A):=\inf\Big\{\ \sum_{n=1}^{\infty}\mu(A_n):\ A\subset\bigcup_{n=1}^{\infty}A_n,\quad A_n\in\mathcal{S},\quad n\in\mathbb{N}\Big\}$$

则 μ^* 为一外测度 (以下称为由 μ 引出的外测度) 且在 $\mathcal{S}$ 上与 μ 一致. 即

$$\mu^*(A)=\mu(A),\qquad\forall\ A\in\mathcal{S}.$$

证明 1) 先证 μ^* 与 μ 在 $\mathcal{S}$ 上一致，若 $A\in\mathcal{S}$, 则由 $A\subset A$ 及 $\mu^*(A)$ 的定义知 $\mu^*(A)\le\mu(A)$(取 $A_1=A$, $A_n=\emptyset$, $n\ge2$). 另一方面，若 $A\subset\bigcup_{n=1}^{\infty}A_n$, $A_n\in\mathcal{S}$, $n\in\mathbb{N}$, 则由引理 10 的 2) 知

$$\mu(A)\le\sum_{n=1}^{\infty}\mu(A_n)$$

由此及 μ^* 的定义知

$$\mu(A)\le\inf\Big\{\sum_{n=1}^{\infty}\mu(A_n):A\subset\bigcup_{n=1}^{\infty}A_n,\{A_n:\ n\in\mathbb{N}\}\subset\mathcal{S}\Big\}=\mu^*(A).$$

2) 再证 μ^* 的不降性.

设 $A\subset B$. 若 $\{\ A_n:\ n\in\mathbb{N}\ \}\subset\mathcal{S}$, 且 $B\subset\bigcup_{n=1}^{\infty}A_n$, 则 $A\subset\bigcup_{n=1}^{\infty}A_n$, 由引理 10 的 2) 及 μ^* 的定义知

$$\begin{aligned}\mu^*(A)&\le\inf\Big\{\ \sum_{n=1}^{\infty}\mu(A_n):B\subset\bigcup_{n=1}^{\infty}A_n,\{\ A_n:n\in\mathbb{N}\ \}\subset\mathcal{S}\ \Big\}\\&=\mu^*(B).\end{aligned}$$

3) 最后证明 μ^* 的次 σ–可加性.

设 $A_n \subset \Omega$, $n \in \mathbb{N}$, 若 $\sum_{n=1}^{\infty} \mu^*(A_n) = \infty$, 则 $\mu^*(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$ 自然成立. 故只考虑 $\sum_{n=1}^{\infty} \mu^*(A_n) < \infty$ 的情形. 此时, 任给 $\varepsilon > 0$, 对任意 $n \in \mathbb{N}$, 由 $\mu^*(A_n)$ 的定义知存在 $\{ A_{nk} : k \in \mathbb{N} \} \subset \mathcal{S}$ 使:

$$A_n \subset \bigcup_{k=1}^{\infty} A_{nk} \text{ 且 } \sum_{n=1}^{\infty} \mu(A_{nk}) \leq \mu^*(A_n) + \frac{\varepsilon}{2^n}$$

成立, 因而 $\bigcup_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} A_{nk}$, 且由 μ^* 的定义及上式知:

$$\begin{aligned} \mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) &\leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mu(A_{nk}) \\ &\leq \sum_{n=1}^{\infty} \Big(\mu^*(A_n) + \frac{\varepsilon}{2^n}\Big) \leq \sum_{n=1}^{\infty} \mu^*(A_n) + \varepsilon. \end{aligned}$$

因 ε 可任意小, 故得 $\mu^*(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$. □

13. 引理 设对任何 $A \subset \mathbb{R}^d$, 令

$$\lambda^*(A) := \inf \Big\{ \sum_{n=1}^{\infty} |I_n| : A \subset \bigcup_{n=1}^{\infty} I_n,\ I_n \text{为} \mathbb{R}^d \text{的开区间} \Big\},$$

其中 $|I_n|$ 表示 I_n 的体积, 则 λ^* 是 $\mathbb{R}^d$ 的一个外测度 (称为 Lebesgue 外测度).

注 由于 $\mathbb{R}^d$ 的所有开区间并不组成一个半集代数, 所以本引理的结论并不是引理 12 的推论, 但证明是类似的. 留给读者作为练习.

14. 定义 设 μ^* 为 Ω 的外测度, $A \subset \Omega$ 称为 μ^*– **可测集**, 如果对任何 $D \subset \Omega$, 有

$$\mu^*(D) = \mu^*(A \cap D) + \mu^*(A^c \cap D).$$

15. 引理 $A\subset\Omega$ 为 μ^* 可测集当且仅当对任何 $D\subset\Omega$ 有

$$\mu^*(D)\geq\mu(A\cap D)+\mu^*(A^c\cap D).$$

证明 显然，只须证条件的充分性. 由 $\mu(\emptyset)=0$ 及 μ^* 的次 σ-可加性，对任何 $D\subset\Omega$, 取 $A_1=A\cap D$, $A_2=A^c\cap D$, $A_n=\emptyset$, $n\geq 3$, 有

$$\mu^*(D)\leq\mu^*(A\cap D)+\mu^*(A^c\cap D).$$

故引理所给条件为充分的. □

16. 定理 设 μ^* 为 Ω 的一个外测度，令

$$\mathcal{A}_{\mu^*}:=\{A:\ A\subset\Omega\text{为 }\mu^*-\text{可测集}\}.$$

则

1) $\mathcal{A}_{\mu^*}$ 为 Ω 的一个 σ 代数；

2) 设 $A_n\in\mathcal{A}_{\mu^*}$, $n\in\mathbb{N}$, 两两不交， $A=\bigcup_{n=1}^{\infty}A_n$, 则

$$\mu^*(A\cap D)=\sum_{n=1}^{\infty}\mu^*(A_n\cap D),\ \forall\ D\subset\Omega;$$

3) μ^* 在 $\mathcal{A}_{\mu^*}$ 上的限制 (仍记作 μ^*) 是 $\mathcal{A}_{\mu^*}$ 上的测度.

证明 (i) 先证 $\mathcal{A}_{\mu^*}$ 是集代数，由 $\mu^*(\emptyset)=0$ 知对任何 $D\subset\Omega$,

$$\mu^*(D)=\mu^*(\Omega\cap D)+\mu^*(\Omega^c\cap D),$$

因而， $\Omega\in\mathcal{A}_{\mu^*}$; 其次，若 $A\in\mathcal{A}_{\mu^*}$, 则由于 μ^* 可测集的定义中 A 和 A^c 的对称性，知 $A^c\in\mathcal{A}_{\mu^*}$; 第三，若 $A,B\in\mathcal{A}_{\mu^*}$, 则对任何 $D\subset\Omega$, 由 μ^* 可测集的定义及次 σ-可加性知

$$\begin{aligned}\mu^*(D)&=\mu^*(A\cap D)+\mu^*(A^c\cap D)\\&=\mu^*(A\cap D)+\mu^*(B\cap(A^c\cap D))+\mu^*(B^c\cap(A^c\cap D))\\&\geq\mu^*((A\cup(A^c\cap B))\cap D)+\mu^*((A\cup B)^c\cap D)\\&=\mu^*((A\cup B)\cap D)+\mu^*((A\cup B)^c\cap D),\end{aligned}$$

故由引理 15 知 $A\cup B\in\mathcal{A}_{\mu^*}$. 因此 $\mathcal{A}_{\mu^*}$ 为集代数.

(ii) 再证 $\mathcal{A}_{\mu^*}$ 是 σ 代数. 为此只需证明: 若 $A_n \in \mathcal{A}_{\mu^*}$, $n \in \mathbb{N}$ 两两不交, 则 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}_{\mu^*}$.

由 μ^*– 可测集的定义知对任何 $D \subset \Omega$,

$$\begin{aligned}\mu^*(D) &= \mu^*(A_1 \cap D) + \mu^*(A_1^c \cap D)\\ &= \mu^*(A_1 \cap D) + \mu^*(A_2 \cap A_1^c \cap D) + \mu^*(A_2^c \cap A_1^c \cap D)\\ &= \mu^*(A_1 \cap D) + \mu^*(A_2 \cap D) + \mu^*((A_1 \cup A_2)^c \cap D)\\ &= \cdots\\ &= \sum_{k=1}^{n} \mu^*(A_k \cap D) + \mu^*\left(\left(\bigcup_{i=1}^{n} A_i\right)^c \cap D\right)\\ &\geq \sum_{k=1}^{n} \mu^*(A_k \cap D) + \mu^*\left(\left(\bigcup_{k=1}^{\infty} A_k\right)^c \cap D\right)\end{aligned}$$

对任何 $n \in \mathbb{N}$ 成立. 令 $n \to \infty$, 并由 μ^* 的次 σ–可加性得

$$\begin{aligned}\mu^*(D) &\geq \sum_{k=1}^{\infty} \mu^*(A_k \cap D) + \mu^*\left(\left(\bigcup_{k=1}^{\infty} A_k\right)^c \cap D\right)\\ &\geq \mu^*\left(\left(\bigcup_{k=1}^{\infty} A_k\right) \cap D\right) + \mu^*\left(\left(\bigcup_{k=1}^{\infty} A_k\right)^c \cap D\right).\end{aligned} \tag{1}$$

由引理 15 即知 $\bigcup_{k=1}^{\infty} A_k \in \mathcal{A}_{\mu^*}$. 因而由 (i) 知 $\mathcal{A}_{\mu^*}$ 为 σ 代数.

(iii) 再证 2) 成立. 设 $A_n \in \mathcal{A}_{\mu^*}$, $n \in \mathbb{N}$, 两两不交, 则由 (ii) 知 $\bigcup_{k=1}^{\infty} A_k \in \mathcal{A}_{\mu^*}$. 且由 (1) 及引理 15 知: 对任何 $D \subset \Omega$ 有

$$\begin{aligned}\mu^*(D) &= \sum_{k=1}^{\infty} \mu^*(A_k \cap D) + \mu^*\left(\left(\bigcup_{k=1}^{\infty} A_k\right)^c \cap D\right)\\ &= \mu^*\left(\left(\bigcup_{k=1}^{\infty} A_k\right) \cap D\right) + \mu^*\left(\left(\bigcup_{k=1}^{\infty} A_k\right)^c \cap D\right).\end{aligned}$$

(若 $\mu^*\left(\left(\bigcup_{k=1}^{\infty} A_k\right)^c \cap D\right) < \infty$, 则由此即得 2), 但此前提未必总成立) 用 $(\bigcup_{k=1}^{\infty} A_k) \cap D$ 代替上式中的 D, 由于 $\mu^*(\emptyset) = 0$, 即得 2).

(iv) 显然 μ^* 非负. 若 $A_n \in \mathcal{A}_{\mu^*}$, $n \in \mathbb{N}$, 两两不交, 则由 2)

(令 $D=\Omega$) 即得

$$\mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big)=\sum_{n=1}^{\infty}\mu^*(A_n).$$

故 μ^* 在 $\mathcal{A}_{\mu^*}$ 上的限制为 $\mathcal{A}_{\mu^*}$ 上的测度.

17. 测度扩张定理的证明

(i) 证明扩张的存在性: 设 μ^* 为由 μ 引出的外测度, 由于 μ^* 与 μ 在 $\mathcal{S}$ 上一致 (引理 12), μ^* 是 $\mathcal{A}_{\mu^*}$ 上的测度, 所以只需证: $\mathcal{S}\subset\mathcal{A}_{\mu^*}$ (由此即知 $\sigma(\mathcal{S})\subset\mathcal{A}_{\mu^*}$!).

设 $A\in\mathcal{S}$, 则由 μ^* 的定义知对任何 $D\subset\Omega$, 当 $\mu^*(D)<\infty$ 时, 对任何 $\varepsilon>0$, 有 $\{A_n\}\subset\mathcal{S}$, 使 $\bigcup_{n=1}^{\infty}A_n\supset D$, 且

$$\sum_{n=1}^{\infty}\mu(A_n)\le\mu^*(D)+\varepsilon. \tag{1}$$

由半集代数的定义及 $A\in\mathcal{S}$ 知存在 $B_k\in\mathcal{S}$, $k=1,\cdots,m$, 两两不交, 使

$$A^c=\Omega\setminus A=\bigcup_{k=1}^{m}B_k.$$

因而

$$\mu(A_n)=\mu(A\cap A_n)+\sum_{k=1}^{m}\mu(A_n\cap B_k). \tag{2}$$

但 $\bigcup_{n=1}^{\infty}(A\cap A_n)\supset A\cap D$, $\bigcup_{n=1}^{\infty}\bigcup_{k=1}^{\infty}(A_n\cap B_k)\supset A^c\cap D$. 于是由 (1), (2) 及 μ^* 的次 σ–可加性 (引理 12)

$$\begin{aligned}\mu^*(D)+\varepsilon&\ge\sum_{n=1}^{\infty}\mu(A\cap A_n)+\sum_{n=1}^{\infty}\sum_{k=1}^{m}\mu(B_k\cap A_n)\\&\ge\mu^*(A\cap D)+\mu^*(A^c\cap D).\end{aligned}$$

由于 ε 可任意小, 所以

$$\mu^*(D)\ge\mu^*(A\cap D)+\mu^*(A^c\cap D).$$

当 $\mu^*(D)=\infty$ 时, 上式显然成立. 因而由引理 15 知 A 是 μ^*–可测集, 即 $A\in\mathcal{A}_{\mu^*}$.

(ii) 设 μ 在 $\mathcal{S}$ 上 σ–有限, 往证扩张的唯一性. 此时由引理 8 知它在 $\mathcal{A}(\mathcal{S})$ 上有唯一扩张, 仍记为 μ, 显然此扩张仍然 σ–有限. 所以存在 $D_n \in \mathcal{A}(\mathcal{S})$, $n \in \mathbb{N}$, 两两不交, $\bigcup_{n=1}^{\infty} D_n = \Omega$, $\mu(D_n) < \infty$, $n \in \mathbb{N}$. 若 μ_1, μ_2 是 μ 在 $\sigma(\mathcal{S})$ 上的两个扩张, 令

$$\Lambda := \left\{ A \in \sigma(\mathcal{S}) : \ \mu_1(A \cap D_n) = \mu_2(A \cap D_n), \ n \in \mathbb{N} \right\}.$$

往证 Λ 为包含 $\mathcal{S}$ 的 λ–系. 事实上, 由于 μ_1, μ_2 在 $\mathcal{A}(\mathcal{S})$ 上与 μ 一致, 故对一切 $A \in \mathcal{S}$, $A \cap D_n \in \mathcal{A}(\mathcal{S})$, $n \in \mathbb{N}$, 且

$$\mu_1(A \bigcap D_n) = \mu(A \bigcap D_n) = \mu_2(A \bigcap D_n),$$

即 $A \in \Lambda$, 故 $\Lambda \supset \mathcal{S}$.

其次, 显然有 $\Omega \in \Lambda$. 若 $A, B \in \Lambda$, $A \subset B$, 则对任何 $n \in \mathbb{N}$, 有

$$\begin{aligned}&\mu_1((B \setminus A) \cap D_n) = \mu_1(B \cap D_n) - \mu_1(A \cap D_n)\\ =&\mu_2(B \cap D_n) - \mu_2(A \cap D_n) = \mu_2((B \setminus A) \cap D_n),\end{aligned}$$

因此, $B \setminus A \in \Lambda$. 若 $A_k \in \Lambda$, $k \in \mathbb{N}$, $A_k \uparrow$, 则令 $A_0 = \varnothing$,

$$\bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} (A_k \setminus A_{k-1}),$$

$A_k \setminus A_{k-1} \in \Lambda$, $k \in \mathbb{N}$, 两两不交, 于是对任何 $n \in \mathbb{N}$, 有

$$\begin{aligned}\mu_1\Big(\Big(\bigcup_{k=1}^{\infty} A_k\Big) \cap D_n\Big) &= \sum_{k=1}^{\infty} \mu_1((A_k \setminus A_{k-1}) \cap D_n)\\ &= \sum_{k=1}^{\infty} \mu_2((A_k \setminus A_{k-1} \cap D_n) = \mu_2\Big(\Big(\bigcup_{k=1}^{\infty} A_k\Big) \cap D_n\Big).\end{aligned}$$

因此 $\bigcup_{k=1}^{\infty} A_k \in \Lambda$. 故 Λ 为一 λ–系. 由 $\mathcal{S}$ 是一 π– 系及定理 1.15 知 $\Lambda = \sigma(\mathcal{S})$, 于是对任何 $A \in \sigma(\mathcal{S})$

$$\mu_1(A) = \sum_{n=1}^{\infty} \mu_1(A \cap D_n) = \sum_{n=1}^{\infty} \mu_2(A \cap D_n) = \mu_2(A).$$

因而扩张唯一. □

18. 推论 由 $\Omega = \mathbb{R}^n$ $(n \in \mathbb{N})$ 上的 Lebesgue 外测度 λ^* (引理 13) 决定的测度空间 $(\mathbb{R}^n, \mathcal{A}_{\lambda^*}, \lambda^*)$ 是 $\mathcal{S}^n$ 上的体积测度 λ, 即

$$\lambda(a,b] := \begin{cases} \prod_{k=1}^{n}(b_k - a_k), & \text{若 } a, b \in \mathbb{R}^n, \\ \lim_{N\to\infty}\prod_{k=1}^{n}(b_k \wedge N - a_k \vee (-N)), & \\ \qquad\qquad \text{若 } \exists\, a_k = -\infty \text{ 或 } b_k = +\infty & \end{cases}$$

的一个扩张, $\mathcal{A}_{\lambda^*}$ 的元称为 n **维 Lebesgue 可测集**, $\mathcal{A}_{\lambda^*}$ 上的测度 λ^* 称为 n 维 **Lebesgue 测度**, λ^* 在 $\mathcal{B}^n = \sigma(\mathcal{S}^n)$ 上的限制称为 n 维 **Borel–Lebesgue 测度** (以后也称为 Lebesgue 测度, 仍记作 λ).

证明 由测度扩张定理知只需证明 λ^* 在 $\mathcal{S}^n$ 上与 λ 一致.

显然, 对任何 $(a,b] \in \mathcal{S}^n$, $\lambda(a,b] \le \lambda^*(a,b]$, 故当 $\lambda(a,b] = \infty$ 时,

$$\lambda(a,b] = \lambda^*(a,b].$$

当 $\lambda(a,b] \in (0,\infty)$ 时, 则 $(a,b]$ 有界, 于是

$$\lambda^*(a,b] \le \lim_{\varepsilon\to 0^+} |(a,b+\varepsilon e)| = \lambda(a,b], \quad e = (1,1,\cdots,1),$$

因而 $\lambda(a,b] = \lambda^*(a,b]$. 当 $\lambda(a,b] = 0$ 时, 则 a, b 的某些坐标相等, 因而 $(a,b] = \varnothing$, 于是由 λ^* 的定义 (引理 13) 立知 $\lambda^*(a,b] = 0$.

故 $\lambda^*(a,b] = 0 = \lambda(a,b]$. □

今后记 λ 在 $\mathcal{A}_{\lambda^*}$ 上的扩张为 λ.

19. 推论 设 F 为 $\mathbb{R}$ 上的右连续增函数, 则在 $\mathcal{B}(\mathbb{R}^1)$ 上有唯一的 σ–有限测度 μ_F 满足:

$$\mu_F(a,b] = F(b) - F(a), \qquad a, b \in \mathbb{R}^1, \quad a \le b.$$

20. 推论 设 F: $\mathbb{R}^n \to \mathbb{R}$ 且满足

1) 对任何 $a \le b$, $a=(a_1,\cdots,a_n)\in\mathbb{R}^n$, $b=(b_1,\cdots,b_n)\in\mathbb{R}^n$,

$$\Delta_{a,b}F := \sum^{n-1}(-1)^k \sum_{1\le i_1<\cdots\le n} F\Big(b-\sum_{\ell=1}^{k}(b_{i_\ell}-a_{i_\ell})e_\ell\Big)\ge 0.$$

其中 e_ℓ 为第 ℓ 个坐标轴上的单位向量.

2) F 关于自变量的每个坐标右连续,

则在 $\sigma(S^n)$ 上有唯一的 σ–有限 μ_F 满足

$$\mu_F(a,b]=\lim_{N\to\infty}\Delta_{a,b}F,\qquad a,b\in\mathbb{R}^n, a\le b$$

μ_F 称为由 F 决定的 **L–S(Lebesgue-Stieltjes)测度**.

习题

1. 设 $(\Omega,\mathcal{F})$ 是可测空间, μ 是 $\mathcal{F}$ 上的可加测度, 且具有次 σ–可加性, 试证 μ 是测度.

2. 设 $(\Omega,\mathcal{F},\mu)$ 是测度空间, 则

1) $\mu(\varnothing)=0$;

2) μ 可加;

3) μ 下方连续: 即对 $\mathcal{F}$ 中任何不降集列 $\{A_n\}$ (即 $A_n\in\mathcal{F}$, 且 $A_n\uparrow$) 都有:

$$\lim_{n\to\infty}\mu(A_n)=\mu\Big(\bigcup_{n=1}^{\infty}A_n\Big);$$

4) μ 上方连续: 即对 $\mathcal{F}$ 中任何不升集列 $\{A_n\}$ (即 $A_n\in\mathcal{F}$, 且 $A_n\downarrow$) 且存在 $m\in\mathbb{N}$, 使 $\mu(A_m)<\infty$, 有

$$\lim_{n\to\infty}\mu(A_n)=\mu\Big(\bigcap_{n=1}^{\infty}A_n\Big).$$

提示: 注意 $A_m\setminus A_n\uparrow$.

3. 设 $(\Omega,\mathcal{F})$ 为可测空间, μ 为 $\mathcal{F}$ 上可加测度, 且满足下述两条件之一:

1) μ 下方连续；

2) $\mu(\Omega) < \infty$ 且对 $\mathscr{F}$ 中任何下降到 $\varnothing$ 的集列 $\{A_n\}$ (即 $A_n \in \mathscr{F}$、且 $A_n \supset A_{n+1}$, $n \in \mathbb{N}$, $\bigcap_{n=1}^{\infty} A_n = \varnothing$) 都有

$$\lim_{n\to\infty} \mu(A_n) = 0 = \mu(\varnothing),$$

则 μ 为 $\mathscr{F}$ 上的测度.

4. 设 $(\Omega, \mathscr{F}, \mu)$ 为测度空间， $\{A_n\}$ 为 $\mathscr{F}$ 中的集序列，记

$$\varliminf_{n\to\infty} A_n := \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k, \qquad \varlimsup_{n\to\infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

试证：

$$\mu\left(\varliminf_{n\to\infty} A_n\right) \le \varliminf_{n\to\infty} \mu(A_n);$$

若存在 $m \in \mathbb{N}$, 使 $\mu\left(\bigcup_{k=m}^{\infty} A_k\right) < \infty$. 则

$$\mu\left(\varlimsup_{n\to\infty} A_n\right) \ge \varlimsup_{n\to\infty} \mu(A_n);$$

若 μ 为有限测度，且 $\varliminf_{n\to\infty} A_n = \varlimsup_{n\to\infty} A_n = A$, 则

$$\mu(A) = \lim_{n\to\infty} \mu(A_n).$$

5. 证明引理 13.

6. 设 $A_n, n \in \mathbb{N}$ 两两不交，且 $\bigcup_{n=1}^{\infty} A_n = \Omega$(称 $\{A_n\}$ 为 Ω 的一个可数划分). 试证： $\mathcal{S} = \{\varnothing\} \cup \{ A_n, \bigcup_{k=n}^{\infty} A_n : n \in \mathbb{N} \}$ 为 Ω 的半集代数.

再设 $q_n \ge 0$、$n \in \mathbb{N}$, 在 $\mathcal{S}$ 上定义 $\mu(A_n) = q_n$、$\mu\left(\bigcup_{k=n}^{\infty} A_k\right) := \sum_{k=n}^{\infty} q_k$、$n \in \mathbb{N}$、$\mu(\varnothing) = 0$. 试具体写出 $\sigma(\mathcal{S})$ 的元及 μ 在 $\sigma(\mathcal{S})$ 上的扩张.

上述测度空间与下列测度空间 $(\Omega, \mathcal{A}, \tilde{\mu})$ 是否相同？此时， $\mathcal{A}$ 为 $A_n, n \in \mathbb{N}$ 的任意并作成的类，

$$\tilde{\mu}\left(\bigcup_{n\in J} A_n\right) = \sum_{n\in J} q_n, \qquad \forall\, J \subset \mathbb{N}.$$

7. 设 $(\Omega,\mathcal{F},\mathbf{P})$ 是概率场，若 $B\in\mathcal{F}$、$\mathbf{P}(B)>0$，则对一切 $A\in\mathcal{F}$，定义

$$\mathbf{P}(A|B)=\frac{\mathbf{P}(A\cap B)}{\mathbf{P}(B)},$$

称为 A **在** B **之下的条件概率**. 试证 $\mathbf{P}(\cdot|B)$ 是 $\mathcal{F}$ 上的概率测度. 因而 $(\Omega,\mathcal{F},\mathbf{P}(\cdot|B))$ 是概率空间.

8. 设 $\{A_k: k\in\mathbb{N}\}$ 是概率空间 $(\Omega,\mathcal{F},\mathbf{P})$ 中随机事件系，试证：

1) 若 $\mathbf{P}(A_n)=1$，$n=1,2,\cdots$，则 $\mathbf{P}(\bigcap_{k=1}^{\infty}A_n)=1$；

2) 若 $\mathbf{P}(A_n)=0$，$n=1,2,\cdots$，则 $\mathbf{P}(\bigcup_{n=1}^{\infty}A_n)=0$.

9. 设 $\mathbf{P}_0$，$\mathbf{P}_1$ 是定义在 $(\Omega,\mathcal{F})$ 上的两个概率测度，若对任意 $\varepsilon>0$，存在 $A_\varepsilon\in\mathcal{F}$，使得 $\mathbf{P}_1(A_\varepsilon)>1-\varepsilon$，而 $\mathbf{P}_0(A_\varepsilon)<\varepsilon$，试证：存在 $A\in\mathcal{F}$，使得 $\mathbf{P}_1(A)=1$，$\mathbf{P}_0(A)=0$.(提示：取 $A=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}A_{1/(2^m)}$).

注 满足上述性质的两个概率测度称为是相互奇异的.

10. 设 $\mathbf{P}',\mathbf{P}$ 是定义在 $(\Omega,\mathcal{F})$ 上的两个概率测度，若对任何使 $\mathbf{P}(A)=0$ 的 $A\in\mathcal{F}$，都有 $\mathbf{P}'(A)=0$，则称 $\mathbf{P}'$ 对 $\mathbf{P}$ 连续，记作 $\mathbf{P}'\ll\mathbf{P}$. 试证：

1) 设 $\mathbf{P}_1,\mathbf{P}_2$ 是定义在 $(\Omega,\mathcal{F})$ 上的两个概率测度，则有 $(\Omega,\mathcal{F})$ 上的一个概率测度 $\mathbf{P}$ 使得 $\mathbf{P}_1\ll\mathbf{P}$ 且 $\mathbf{P}_2\ll\mathbf{P}$.

2) 将 1) 推广至无穷多个概率测度 $\mathbf{P}_1,\mathbf{P}_2,\cdots$ 的情形.

11. 设 f 是增函数且存在实数 A 与 B 使得对任何 x：$A\leq f(x)\leq B$. 证明对每个 $\varepsilon>0$，大小超过 ε 的跳跃数最多为 $(B-A)/\varepsilon$. 由此证明：任何不降函数 f 的不连续点集合最多可数. (提示：首先就 f 有界的情形来证明，然后考虑一般情况.)

12. 设 f 是 $(-\infty,+\infty)$ 上的一个任意函数，L 是所有这种 x 的集：f 在 x 处右连续但不左连续，证明 L 是一可数集. [提

示：考虑 $L \cap M_n$，其中 $M_n = \{ x | O(f;x) > 1/n \}$. $O(f;x) = \inf\limits_{\varepsilon > 0} \sup\limits_{t \in (x-\varepsilon, x+\varepsilon)} |f(t) - f(x)|$.

13. 设 f 是 D 上增函数，D 在 $(-\infty, +\infty)$ 中稠密，在 $\mathbb{R}$ 上如下定义 $\tilde{f}$：

$$\tilde{f}(x) = \inf_{x < t \in D} f(t), \qquad \forall x \in \mathbb{R}.$$

证明：f 在 D 上的一致连续性一定蕴涵 $\tilde{f}$ 在 $\mathbb{R}$ 上的一致连续性；并举一反例说明 f 在 D 上的连续性并不蕴涵 $\tilde{f}$ 在 $\mathbb{R}$ 上的连续性.

*14. 任给 $\mathbb{R}$ 上的广义实值函数 f，存在可数集 D 具有如下性质，对于每个 t，存在 $t_n \in D$，$n \in \mathbb{N}$，$t_n \to t$ $(n \to \infty)$，使得 $f(t) = \lim_{n\to\infty} f(t_n)$. 如果 "$t_n \to t$"，用 "$t_n \downarrow t$" 或 "$t_n \uparrow t$" 来代替，上述论断仍成立.［这是随机过程"可分性"的关键，考虑图形 $(t, f(t))$，并引入度量

$$d((x, f(x)), (y, f(y))) := |x - y| + \left| \frac{f(x)}{1 + |f(x)|} - \frac{f(y)}{1 + |f(y)|} \right|,$$

先证可分距离空间的子空间仍然可分，再应用这个结果证明在上述距离下，$\{ (x, f(x)) : x \in \mathbb{R} \}$ 是可分距离空间，由此即可得第一结论.］

15. 计算下列各 Borel 集的 Lebesgue 测度：

1) $[0, 1]$ 中的无理点集；2) Cantor 集；3) Sierpinski 海绵；4) 圆周；5) 开圆 $B(0, 1)$；6) $[a, b]$ 上的连续函数 $f(x)$，$x \in [a, b]$ 的图 $\{ (x, f(x)) : x \in [a, b] \}$.

16. 若 $\mu^*(A) = 0$，则 $\mu^*(A \cup B) = \mu^*(B)$.

17. 若 $\mathbb{R}^n$ 的有界闭区间 $[a, b]$ 至少有一边长为 0 (即至少有一 k, $1 \le k \le n$, 使 $a_k = b_k$) 则 $\lambda^*[a, b] = 0$.

18. 设 μ,ν 是可测空间 $(\Omega,\mathcal{F})$ 上的两个有限测度， $\mathcal{C}$ 为 π–系， $\Omega\in\mathcal{C}$ 且 $\sigma(\mathcal{C})=\mathcal{F}$. 若 μ,ν 在 $\mathcal{C}$ 上一致，则 μ,ν 在 $\mathcal{F}$ 上一致.

19. 设 $\Omega=\{a,b,c,d\}$, $\mathcal{C}=\{\{a,b\},\{c,d\},\{a,c\},\{b,d\},\Omega,\emptyset\}$，令：

$$\mu_1(\{a\})=\mu_1(\{d\})=\mu_2(\{b\})=\mu_2(\{c\})=1;$$

$$\mu_1(\{b\})=\mu_1(\{c\})=\mu_2(\{a\})=\mu_2(\{d\})=2;$$

$$\mu_i(\{x,y\})=\mu_i(\{x\})+\mu_i(\{y\}),\qquad i=1,2,\ \{x,y\}\in\mathcal{C}.$$

$$\mu_i(\emptyset)=0,\qquad \mu_i(\Omega)=6,\quad i=1,2.$$

试证： $\mathcal{C}$ 不是半集代数，在 $\mathcal{C}$ 上 $\mu_1=\mu_2$ 且都 σ–可加. 但在 $\sigma(\mathcal{C})$ 上 $\mu_1\neq\mu_2$. 这个例子说明了什么问题？

20. 如果测度扩张定理中的 "σ–有限" 条件去掉，则扩张的唯一性未必成立. 例如 $\Omega=\mathbb{R}$，在 $\mathcal{S}$ 上定义 $\mu_1(A)=$ 集 A 的势，$A\in\mathcal{B}$; $\mu_2=2\mu_1$. 试证： μ_1,μ_2 都是 $\mathcal{B}$ 上的测度，不是 σ–有限的，但在 $\mathcal{S}$ 上 $\mu_1=\mu_2$. 而在 $\sigma(\mathcal{S})=\mathcal{B}$ 上， $\mu_1\neq\mu_2$.

21. 设 $(\Omega,\mathcal{F},\mu)$ 为一可测空间， Δ 为 Ω 的任一子集，令

$$\mu_\Delta(A):=\mu^*(\Delta\cap A).$$

试证：$(\Delta,\Delta\cap\mathcal{F},\mu_\Delta)$ 为一以 Δ 为空间的测度空间. 若 $0<\mu^*(\Delta)<\infty$，则 $\left(\Delta,\Delta\cap\mathcal{F},\dfrac{\mu_\Delta}{\mu^*(\Delta)}\right)$ 为一以 Δ 为样本空间的概率空间.

*22. 设 (E,ρ) 为一完全可分距离空间， $K_n\subset E$, $n\in\mathbb{N}$, 为 E 的一个上升的紧集列， $\{S_n: n\in\mathbb{N}\}$ 是作为 E 的可数拓扑基的开球列， $\mathcal{F}$ 为 $S_m\cap K_n$, $m,n\in\mathbb{N}$, 的一切有限并作成的集类， $\nu:\mathcal{F}\mapsto\mathbb{R}_+$，具有性质：

(i) $F_i\in\mathcal{F}$, $i=1,2$, $F_1\subset F_2\implies\nu(F_1)\le\nu(F_2)$;

(ii) $F_1,F_2\in\mathcal{F}\implies\nu(F_1\cup F_2)\le\nu(F_1)+\nu(F_2)$;

(iii) $F_1, F_2 \in \mathcal{F}$, $F_1 \cap F_2 = \emptyset \implies \nu(F_1 \cup F_2) = \nu(F_1) + \nu(F_2)$.
对于 E 的开集 O 及任何 $A \subset E$. 定义

$$\lambda(O) := \sup\{\nu(F): F \subset O, F \in \mathcal{F}\},$$

$$\lambda^*(A) := \inf\{\lambda(G): A \subset G, G\text{为开集}\}.$$

则 λ^* 为 E 上的一个外测度.

(此题来自 [LR] 中 Prohorov 定理的证明过程.)

§3. 测度空间的一些性质

今后如无特别声明, 谈到测度时总是指某一给定可测空间的 σ–代数上的测度.

一. 测度的运算性质.

2.1 中定义的可测空间 $(\Omega, \mathcal{F})$ 上的测度 μ 是 $\mathcal{F}$ 上的广义实值函数, 满足:

1) 对任何 $A \in \mathcal{F}$. $\mu(A) \geq 0$. 且至少存在 $A \in \mathcal{F}$. 使 $\mu(A) < \infty$;

2) 若 $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, 两两不交, 则

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n).$$

若还满足

3) $\mu(\Omega) = 1$; 则称 μ 为 $(\Omega, \mathcal{F})$ (或 $\mathcal{F}$) 上的概率测度.

1. 由上述定义, 易得

4) $\mu(\emptyset) = 0$;

5) $A, B \in \mathcal{F} \implies \mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$
若还有 $\mu(A) < \infty$, $\mu(B) < \infty$. 则

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B).$$

一般地，若 $\mu(A_k)<\infty$, $k=1,2,\cdots,n$, 则

$$\mu(A_1\cup\cdots\cup A_n)=\sum_{k=1}^{n}(-1)^{k-1}\sum_{1\le i_1<\cdots<i_k\le n}\mu(A_{i_1}\cdots A_{i_k})$$

6) 若 A, $B\in\mathcal{F}$, $A\subset B$, 则 $\mu(B)=\mu(A)+\mu(B\setminus A)$, 且 $\mu(A)\le\mu(B)$; 若还有 $\mu(A)<\infty$, 则

$$\mu(B\setminus A)=\mu(B)-\mu(A);$$

若 μ 是概率测度，则

$$\mu(A^c)=1-\mu(A);$$

$$\mu(A)\le 1;$$

7) 次 σ–可加性：若 $A_n\in\mathcal{F}$, $n\in\mathbb{N}$, 则

$$\mu\Big(\bigcup_{n=1}^{\infty}A_n\Big)\le\sum_{n=1}^{\infty}\mu(A_n).$$

2. 定理 设 $(\Omega,\mathcal{F},\mu)$ 是测度空间, 则 μ 下方连续:即若 $A_n\in\mathcal{F}$, $n\in\mathbb{N}$, $A_n\uparrow$, 则

$$\lim_{n\to\infty}\mu(A_n)=\mu\Big(\bigcup_{n=1}^{\infty}A_n\Big).$$

且 μ 上方连续：即若 $A_n\in\mathcal{F}$, $n\in\mathbb{N}$, $A\downarrow$ 且存在 $m\in\mathbb{N}$, 使 $\mu(A_m)<\infty$, 则

$$\lim_{n\to\infty}\mu(A_n)=\mu\Big(\bigcap_{n=1}^{\infty}A_n\Big).$$

证明 若 $A_n\in\mathcal{F}$, $n\in\mathbb{N}$, $A_n\uparrow$, 则 (令 $A_0=\varnothing$)

$$\begin{aligned}\lim_{n\to\infty}\mu(A_n)&=\lim_{n\to\infty}\sum_{k=1}^{\infty}\mu(A_k\setminus A_{k-1})=\sum_{k=1}^{\infty}\mu(A_k\setminus A_{k-1})\\&=\mu\Big(\bigcup_{k=1}^{\infty}(A_k\setminus A_{k-1})\Big)=\mu\Big(\bigcup_{k=1}^{\infty}A_n\Big).\end{aligned}$$

若 $A_n\in\mathcal{F}$, $n\in\mathbb{N}$, $A_n\downarrow$ 且存在 $m\in\mathbb{N}$, 使 $\mu(A_m)<\infty$, 则

$$A_n=\Big(\bigcup_{k=n}^{\infty}(A_k\setminus A_{k+1})\Big)\cup\Big(\bigcap_{k=1}^{\infty}A_k\Big)$$

且当 $n \geq m$ 时，$\mu(A_n) < \infty$，因而

$$\mu(A_n) = \sum_{k=n}^{\infty} \mu(A_k \setminus A_{k+1}) + \mu\Big(\bigcap_{k=1}^{\infty} A_k\Big)$$

中的级数收敛，当 $n \to \infty$ 时，极限为 0，故

$$\lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \sum_{k=n}^{\infty} \mu(A_k \setminus A_{k+1}) + \mu\Big(\bigcap_{k=1}^{\infty} A_k\Big)$$
$$= \mu\Big(\bigcap_{k=1}^{\infty} A_k\Big).$$

(也可用 de Morgan 法则归结到前一部分)．□

3. 定理 设 μ 为 Ω 的集代数 $\mathcal{A}$ 上的可加测度，若 μ 还满足下列两条件之一：

1) μ 下方连续，即对任何满足 $\{A_n\} \subset \mathcal{A}$，$A_n \uparrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ 的集列，都有

$$\lim_{n\to\infty} \mu(A_n) = \mu(\bigcup_{n=1}^{\infty} A_n);$$

2) $\mu(\Omega) < \infty$ 且在 $\emptyset$ 处上方连续，即对任何满足 $\{A_n\} \subset \mathcal{A}$、$A_n \downarrow \bigcap_{n=1}^{\infty} A_n = \emptyset$ 的集列，都有

$$\lim_{n\to\infty} \mu(A_n) = 0 = \mu(\emptyset),$$

则 μ 为 $\mathcal{A}$ 上的测度．

证明 只需证 μ 在 $\mathcal{A}$ 上具有 σ–可加性．事实上，由 $\mathcal{A}$ 是集代数知 μ 可加蕴含有限可加．对任何 $B_n \in \mathcal{A}$，$n \in \mathbb{N}$，两两不交，且 $\bigcup_{n=1}^{\infty} B_n \in \mathcal{A}$．

若 1) 成立，令 $A_n := \bigcup_{k=1}^{n} B_k \in \mathcal{A}$，则 $A_n \uparrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$，因而

$$\mu\Big(\bigcup_{k=1}^{\infty} B_k\Big) = \mu\Big(\bigcup_{k=1}^{\infty} A_n\Big) = \lim_{n\to\infty} \mu(A_n)$$
$$\xlongequal{\text{有限可加}} \lim_{n\to\infty} \sum_{k=1}^{n} \mu(B_k) = \sum_{k=1}^{\infty} \mu(B_k).$$

若 2) 成立，令 $A_n = \bigcup_{k=n+1}^{\infty} B_k$, 则 $A_n \in \mathcal{A}$ 且 $A_n \downarrow \varnothing$ (为什么？), 因而

$$\lim_{n\to\infty} \mu\Big(\bigcup_{k=n+1}^{\infty} B_k\Big) = \lim_{n\to\infty} \mu(A_n) = 0.$$

由有限可加性知

$$\begin{aligned}\mu\Big(\bigcup_{k=1}^{\infty} B_k\Big) &= \mu\Big(\bigcup_{k=1}^{n} B_k\Big) + \mu\Big(\bigcup_{k=n+1}^{\infty} B_k\Big)\\ &= \sum_{k=1}^{n} \mu(B_k) + \mu\Big(\bigcup_{k=n+1}^{\infty} B_k\Big).\end{aligned}$$

令 $n\to\infty$ 即得

$$\mu\Big(\bigcup_{k=1}^{\infty} B_k\Big) = \sum_{k=1}^{\infty} \mu(B_k). \quad \square$$

二、测度的完全化.

4. 定义 设 $(\Omega, \mathcal{F}, \mu)$ 是一测度空间. 若对 Ω 的子集 B 存在 $A \in \mathcal{F}$ 使 $A \supset B$ 且 $\mu(A) = 0$, 则称 B 为一 μ–**零集**. 若每一 μ–零集都属于 $\mathcal{F}$, 则称 $(\Omega, \mathcal{F}, \mu)$ 为 **完全测度空间**, 称 μ 为 **完全测度**.

5. 定理 1) 设 $(\Omega, \mathcal{F}, \mu)$ 是一测度空间. 令

$$\begin{aligned}\overline{\mathcal{F}} :&= \{ A\Delta N:\ A \in \mathcal{F},\ N \text{ 为 } \mu-\text{零集} \}\\ &= \{ A\cup N:\ A \in \mathcal{F},\ N \text{ 为 } \mu-\text{零集} \}\end{aligned}$$

$$\overline{\mu}(A\Delta N) := \mu(A),\ A \in \mathcal{F},\ N \text{ 为 } \mu-\text{零集}.$$

则 $(\Omega, \overline{\mathcal{F}}, \overline{\mu})$ 为一完全测度空间, 称它为 $(\Omega, \mathcal{F}, \mu)$ 的 **完全化**.

2) 设 μ 为半集代数 $\mathcal{S}$ 上的 σ–有限测度, μ^* 是由 μ 引出的外测度, $\mathcal{A}^*$ 是一切 μ^*–可测集组成的 σ 代数, 则测度空间 $(\Omega, \mathcal{A}^*, \mu^*)$ 是 $(\Omega, \sigma(\mathcal{S}), \mu)$ 的完全化 (此处将 μ 在 $\sigma(\mathcal{S})$ 上的扩张仍记作 μ).

因而此时 μ 在 $\mathcal{A}^*$ 上的扩张是唯一的.

6. 推论 $\mathbb{R}^n$ 上的 Lebesgue 测度是 Borel 测度的完全化.

定理 5. 1) 可以直接验证, 而定理 5 即为 [YWL] 第一章 §5.3 的定理 5-8. 由于篇幅原因此处不再证明了. 推论 6 是定理 5. 2) 及推论 2.17 的直接结果.

完全化的好处在于: 假设某个依赖于 ω 的性质在某个零测集 N 之外成立, 则使此性质不成立的 ω 的集就是 N 的一个子集, 一般来说, 它不一定属于 $\mathcal{F}$, 但它属于 $\overline{\mathcal{F}}$, 且它的 $\overline{\mu}$ 测度为零. 有了这个例外集的可测性有时是方便的.

三. 测度的逼近.

由测度扩张定理的证明知: 由半集代数 $\mathcal{S}$ 上的测度 μ 引出外测度 μ^*, 再定出 μ^* 可测集类 $\mathcal{A}_{\mu^*}$ 及测度 $\mu^*|_{\mathcal{A}_{\mu^*}}$. 但是 $\mathcal{A}_{\mu^*}$ 的一般元素及其外测度都不太容易把握. 因此有时考虑它们的逼近问题, 下面是这方面的一些结果.

7. 命题 设 μ 是半集代数 $\mathcal{S}$ 上的测度, $\mathcal{A}(\mathcal{S})$ 是 $\mathcal{S}$ 生成的集代数 (因而由引理 2.7 知 μ 在它上面的扩张是相当具体的), μ^* 是由 μ 引出的外测度, 则对任何 $A \in \mathcal{A}_{\mu^*}$, $\mu^*(A) < \infty$ 及对任何 $\varepsilon > 0$, 存在 $A_\varepsilon \in \mathcal{A}(\mathcal{S})$ 使

$$\mu^*(A\Delta A_\varepsilon) = \mu^*(\{\ \omega:\ |I_A(\omega) - I_{A_\varepsilon}(\omega)| = 1\ \}) < \varepsilon.$$

证明 由 μ^* 的定义知存在 $B_n \in \mathcal{S}, n \in \mathbb{N}$, 使 $A \subset \bigcup_{n=1}^\infty B_n$ 且

$$\mu^*(A) \le \sum_{n=1}^\infty \mu(B_n) < \mu^*(A) + \frac{\varepsilon}{2}.$$

因 $\mu^*(A) < \infty$, 所以 $\sum_{n=1}^\infty \mu(B_n)$ 收敛, 故存在 $n_0 \in \mathbb{N}$, 使

$$\sum_{n=n_0+1}^\infty \mu(B_n) < \frac{\varepsilon}{2}.$$

令 $A_\varepsilon := \bigcup_{n=1}^{n_0} B_n$, $B_\varepsilon := \bigcup_{n=n_0+1}^{\infty} B_n$. 则

$$A \setminus A_\varepsilon \subset B_\varepsilon, \quad A_\varepsilon \setminus A \subset (A_\varepsilon \cup B_\varepsilon) \setminus A,$$

$$\begin{aligned}\mu^*(A \Delta A_\varepsilon) &\le \mu^*(A \setminus A_\varepsilon) + \mu^*(A_\varepsilon \setminus A)\\ &\le \mu^*(B_\varepsilon) + \mu^*\Big(\Big(\bigcup_{n=1}^{\infty} B_n\Big) \setminus A\Big)\\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \quad \Box\end{aligned}$$

对于 $\mathbb{R}^n$ 上的 Lebesgue 测度，还可以有如下进一步的

8. 命题 设 A 为 $\mathbb{R}^n$ 的一个 Lebesgue 可测子集，则

$$\begin{aligned}\lambda(A) &= \inf\{\ \lambda(G):\ A \subset G \subset \mathbb{R}^n,\ G\text{是开集}\ \},\\ &= \sup\{\ \lambda(K):\ A \supset K,\ K\text{是紧集}\ \}.\end{aligned}$$

因而若 $\lambda(A) < \infty$, 则对任何 $\varepsilon > 0$, 存在一个开集 G 及一个紧集 K (因而是有界闭集) 使

$$K \subset A \subset G\ \text{且}\lambda(G) - \varepsilon \le \lambda(A) \le \lambda(K) + \varepsilon.$$

证明 由定理 2.2.11 知开集是可列个左开右闭区间之并，所以是 Borel 集，故它们都是 Lebesgue 可测集. 为证定理的前一部分，由 λ 的不降性知，只须证

(1) $\lambda(A) \ge \inf\{\ \lambda(G):\ A \subset G \subset \mathbb{R}^n,\ G\ \text{是开集}\ \}$

(2) $\lambda(A) \le \sup\{\ \lambda(K):\ A \supset K,\ K\ \text{是紧集}\ \}$.

若 $\lambda(A) = \infty$, 则 (1) 显然成立，设 $\lambda(A) < \infty$, $\varepsilon > 0$ 任意，则由 Lebesgue 测度 λ 的定义知存在开区间列 I_k, $k \in \mathbb{N}$, 使

$$A \subset \bigcup_{k=1}^{\infty} I_k, \qquad \sum_{k=1}^{\infty} |I_k| < \lambda(A) + \varepsilon.$$

令 $G := \bigcup_{k=1}^{\infty} I_k$, 则 $A \subset G$ 且由 λ 的次 σ–可加性

$$\lambda(G) \le \sum_{k=1}^{\infty} |I_k| \le \lambda(A) + \varepsilon.$$

故 (1) 成立.

对于 (2)，首先讨论 A 为有界集的情形 (主要想法是利用余集). 此时必有一有界闭集 $C \supset A$. 于是由前一部分证明存在 G 为开集，使

$$G \supset C \setminus A, \qquad \lambda(G) \leq \lambda(C \setminus A) + \varepsilon.$$

则 $K := C \setminus G$ 为有界闭集，因而是紧集，且 $K \subset A$

$$\lambda(K) = \lambda(C) - \lambda(C \cap G) \geq \lambda(C) - \lambda(C \setminus A) - \varepsilon = \lambda(A) - \varepsilon.$$

由于 $\varepsilon > 0$ 任意，所以 (2) 成立.

现在对无界 Lebesgue 可测集 A 来证明 (2). 为此只需证明：对任何 $b < \lambda(A)$ 都有一紧集 $K \subset A$ 使 $\lambda(K) > b$ 成立. 由 λ 的下连续性知

$$\lambda(A) = \lim_{m \to \infty} \lambda(A \cap [-me, me]), \qquad e = (1, 1, \cdots, 1).$$

于是，存在 $m_0 \in \mathbb{N}$ 使

$$\lambda(A \cap [-m_0 e, m_0 e]) > b,$$

由前一段证明知存在 $K \subset A$, K 为紧集使 $\lambda(K) > b$. □

四. Lebesgue 测度还有一些其他值得注意的性质. 由于篇幅关系，我们在下面不加证明地加以介绍.

9. 命题 设 μ 是 $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ 上的测度，且满足：

1) 对任何有界 Borel 集 A, $\mu(A) < \infty$;

2) μ 平移不变，即对任何 $x \in \mathbb{R}^n$, $A \in \mathcal{B}(\mathbb{R}^n)$

$$\mu(A) = \mu(x + A)$$

其中 $x + A := \{ y \in \mathbb{R}^n : y = x + a, a \in A \}$，则有一常数 $c \geq 0$，使对任何 $A \in \mathcal{B}(\mathbb{R}^n)$, $\mu(A) = c\lambda(A)$.

10. 命题 设 A 是 $\mathbb{R}$ 的 Lebesgue 可测子集且 $\lambda(A) > 0$. 则存在 $\varepsilon > 0$ 使

$$(-\varepsilon, \varepsilon) \subset \mathrm{diff}(A) := \{\, x - y : \ x, y \in A \,\}.$$

11. 命题 $\mathbb{R}$ 上有不 Lebesgue 可测的子集存在.

五. Hausdorff 维数及测度介绍.

我们已经证明 Cantor 集是不可数集. 而且容易证明:

12. 命题 Cantor 集 P_0 的 Lebesgue 测度是 0.

证明 G_0 的 Lebesgue 测度为

$$\frac{1}{3} + \frac{1}{3}\left(\frac{2}{3}\right) + \frac{1}{3}\left(\frac{2}{3}\right)^2 + \cdots + \frac{1}{3}\left(\frac{2}{3}\right)^n + \cdots = \frac{1}{3} \cdot \frac{1}{1 - \frac{2}{3}} = 1.$$

所以 $\lambda(P_0) = 1 - \lambda(G_0) = 0$. □

给定 $s > 0$, $\eta > 0$. 对任何 $A \subset \mathbb{R}^n$, 定义

$$\delta(F) := \sup\{d_2(x, y) : x, y \in F\} \text{ 表示集} F \text{ 的直径},$$

$$\mathcal{H}^s_\eta(A) := \inf\left\{ \sum_{k=1}^{\infty} [\delta(U_k)]^s : \ A \subset \bigcup_{k=1}^{\infty} U_k, U_k \subset \mathbb{R}^n, 0 < \delta(U_k) \le \eta \right\}$$

当 $\eta \downarrow$ 时, 上式右边的集减小, $\mathcal{H}^s_\eta(A) \uparrow$. 于是

$$\mathcal{H}^s(A) := \lim_{\eta \to 0} \mathcal{H}^s_\eta(A)$$

存在, 可以证明 $\mathcal{H}^s$ 是 $\mathbb{R}^n$ 的一个外测度, 而且在 $\mathcal{B}^n$ 上的限制是一个测度. (完全仿照推论 2.12 的证法可证: 对任何 $\eta > 0$. $\mathcal{H}^s_\eta$ 是一外测度, 因而 $\mathcal{H}^s$ 是一外测度.)

引出 $\mathcal{H}^s_\eta(A)$ 的 "想法":

如果 U_k 是 d 维球, 则 $(\delta(U_k))^d$ 与球 U_k 的体积只相差一个常数倍, 因而当 s 为整数时, $\mathcal{H}^s_\eta(A)$ 就是 A 的 s 维的体积 (不管 η 多大!).

另外，对任何 $A \subset \mathbb{R}^n$, $\eta < 1$ 容易看出 $\mathcal{H}_\eta^s(A)$ 对 s 是不增的，因此 $\mathcal{H}^s(A)$ 对 s 也是不增的，更进一步，当 $t > s$ 时，

$$\sum_k [\delta(U_k)]^t \le \eta^{t-s} \sum_k [\delta(U_k)]^s.$$

因而

$$\mathcal{H}_\eta^t(A) \le \eta^{t-s} \mathcal{H}_\eta^s(A)$$

于是 (令 $\eta \to 0$)

$$\mathcal{H}^s(A) < \infty, \qquad t > s \implies \mathcal{H}^t(A) = 0$$

由此即知存在 $\dim_H A \in [0, n]$ 使

$$\mathcal{H}^s(A) = \begin{cases} 0, & s > \dim_H(A); \\ \infty, & s < \dim_H(A). \end{cases}$$

而称 $\dim_H(A)$ 为 A 的 Hausdorff 维数.

可以算出

$$\dim_H P_0 = \log 2 / \log 3.$$

Hausdorff 维数与 Hausdorff 测度问题是近些年发展起来的分形几何 (Fractal Geometry) 的一个基本概念. 关于分形几何可参看 [Fa1], [Fa2].

一个富有启发性的算法如下：

$$\begin{aligned} \mathcal{H}^s(P_0) &= \mathcal{H}^s(P_0 \cap [0, 1/3]) + \mathcal{H}^s(P_0 \cap [2/3, 1]) \\ & \left(\frac{1}{3}\right)^s \mathcal{H}^s(P_0) + \left(\frac{1}{3}\right)^s \mathcal{H}^s(P_0) \end{aligned}$$

假设在 $s = \dim_H P_0$ 下有 $0 < \mathcal{H}^s(P_0) < \infty$ (此式需要证明，是一个大胆而合理的假设，此事证明常有实质性困难，有时不成立!) 于是得到 $1 = 2\left(\frac{1}{3}\right)^s$. 因而 $\dim_H P_0 = s = \log 2 / \log 3$.

习题

1. 设 $\mathbf{P}_0, \mathbf{P}_1$ 是 $(\Omega, \mathcal{F})$ 上的两个概率测度. 且对任何 $\varepsilon > 0$, 存在 $A_\varepsilon \in \mathcal{F}$ 使 $\mathbf{P}_0(A_\varepsilon) < \varepsilon$, $\mathbf{P}_1(A_\varepsilon) > 1 - \varepsilon$. 试证: 存在 $A \in \mathcal{F}$ 使 $\mathbf{P}_1(A) = 1$, $\mathbf{P}_0(A) = 0$ (提示: 取 $A = \overline{\lim}_{n\to\infty} A_{2^{-n}}$) (称具有上述性质的两个测度是相互奇异的).

2. 设 $(\Omega, \mathcal{F}, \mathbf{P})$ 是概率空间, 若 $\{ A_n : n \in \mathbb{N} \} \subset \mathcal{F}$ 且 $\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty$. 则

$$\mathbf{P}(\{ A_n \text{ i.o. } \}) = 0$$

(其中 $\{ A_n \text{ i.o.} \}$ 表示有无穷个 A_n 发生的事件, 试证

$$\{ A_n \text{ i.o.} \} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_n).$$

此结论是著名的 Borel -Cantelli 引理.

3. 设 $\mathcal{A}$ 为 Ω 上的集代数.

1) 设 μ_1, μ_2 是 $\sigma(\mathcal{A})$ 上的测度且在 $\mathcal{A}$ 上 σ–有限, 若 $\mu_i(A) < \infty$, $i = 1, 2$, 则对任何 $\varepsilon > 0$, 存在 $A_\varepsilon \in \mathcal{A}$ 使 $\mu_i(A \Delta A_\varepsilon) < \varepsilon$, $i = 1, 2$.

2) 试将 1) 推广至可数个测度的情形.

4. 证明 $A \subset \mathbb{R}^n$ Lebesgue 可测的充要条件是对任何 $I \subset \mathbb{R}^n$ 开区间, 有

$$\lambda^*(I) = \lambda^*(A \cap I) + \lambda^*(A^c \cap I).$$

5. 若 I 为 $\mathbb{R}^n$ 的一个有界开区间, 试证:
$B \subset I$ 为 Lebesgue 可测集的充要条件是

$$\lambda^*(I) = \lambda^*(B) + \lambda^*(I \cap B^c).$$

如果定义 $\lambda_*(B) = \sup\{ \lambda(F) : F \subset B, F \text{ 为闭集 } \}$(称为 B 的内测度), 试证: 上述条件等价于 $\lambda^*(B) = \lambda_*(B)$.

6. 证明对任一集 $E \subset \mathbb{R}$, 有 G_δ 型集 (可表示为可数个开集的并) A, 使 $\lambda(A) = \lambda^*(E)$.

7. $E \subset \mathbb{R}$ 是 Lebesgue 可测集 $\iff$ $E = A \setminus H$, 其中 A 为 G_δ 型集, $\lambda(H) = 0$.

8. $E \subset \mathbb{R}$ 是 Lebesgue 可测集 $\iff$ $E = B \cup H$, 其中 B 为 F_σ 型集 (可表示为可数个闭集的交), 而 $\lambda(H) = 0$.

9. 设 (E, d) 为距离空间, 定义:

i) 对一切 E 的子集 A, B. 称 $d(A, B) := \inf\{d(x, y) : x \in A, y \in B\}$, 为集合 A, B 的距离, 如果 $d(A, B) > 0$, 称 A, B 是正分离的;

ii) 对任何 E 的子集 A. 称 $\delta(A) := \sup\{d(x, y) : x, y \in A\}$ 为集合 A 的直径;

iii) 如果 μ 是 E 上的外测度, 对任何正分离的集合 A, B 有 $\mu(A \bigcup B) = \mu(A) + \mu(B)$, 则称 μ 为距离外测度;

iv) 若对 E 的任何子集 A, $s > 0$, $\varepsilon > 0$, 令

$$\mathcal{H}_\varepsilon^s(A) := \inf\{\sum_{i=1}^{\infty} \delta^s(U_i) : A \subset \bigcup_{i=1}^{\infty} U_i, 0 < \delta(U_i) \leqslant \varepsilon\},$$

$$\mathcal{H}^s(A) := \lim_{\varepsilon \to 0} \mathcal{H}_\varepsilon^s(A) = \sup_{\varepsilon > 0} \mathcal{H}_\varepsilon^s(A).$$

称 $\mathcal{H}^s(A)$ 为集合 A 的 Hausdorff 测度.

试证:

1) **引理** 若 μ 为 (E, d) 上的距离外测度, $\{A_n\} \subset E$ 为不降集列, $A = \lim\limits_{n \to \infty} A_n$, 且对任何 $n \in \mathbb{N}$, $d(A_n, A \setminus A_{n+1}) > 0$, 则 $\mu(A) = \lim\limits_{n \to \infty} \mu(A_n)$;

2) **定理** 若 μ 为 (E, d) 上的距离外测度, 则 E 的一切 Borel 子集是 μ-可测集.

3) **定理** $\mathcal{H}^s$ 是 (E, d) 上的距离外测度.

第四章 可测函数与随机变量

§1. 可测函数与分布

1. 定义 设 $(\Omega, \mathcal{F}, \mu)$ 是一测度空间，若函数 $f: \Delta(\in \mathcal{F}) \mapsto \overline{\mathbb{R}}$ 使对一切 $B \in \overline{\mathcal{B}}$，有

$$f^{-1}(B) := \{\omega \in \Delta: f(\omega) \in B\} \in \Delta \cap \mathcal{F},$$

其中 $\Delta \cap \mathcal{F} := \{\Delta \cap F: F \in \mathcal{F}\}$，则称 f 为定义在 Δ **上的可测函数**；当 μ 为概率测度时，可测函数称为 Δ **上的随机变量** (简记作 r.v.)。若 f 取值于 $\mathbb{R}$ 时，则称 f 为 Δ 上的 **有限实值可测函数**(相应地，**有限值 r.v.**)。

一个 Δ **上的有限复值可测函数**(相应地，**复值 r.v.**) 是一个定义在 $\Delta \in \mathcal{F}$ 而取值于复平面的函数，且其实部和虚部都是有限实值可测函数。以后复值函数都可按此定义化成实的情形讨论。

若 $\Delta = \Omega$，则相应的名词前的 "Δ 上的" 略去，并有时简称为 $\mathcal{F}$–可测函数。我们知道 $(\Delta, \Delta \cap \mathcal{F}, \mu)$ 也是一个测度空间，所以如果有必要，可以将 "Δ 上的" 化为定义在整个空间上的问题来讨论。例如对于 r.v. X，要想考虑有限值的情况，就可考虑

$$\Delta_0 := \{\omega \in \Omega: |X(\omega)| < \infty\}$$

上的 r.v.

更一般地，设 $(\Omega, \mathcal{F})$、$(E, \mathcal{E})$ 为两个可测空间，映射 $f: \Omega \mapsto E$ 满足：对一切 $B \in \mathcal{E}$，有 $f^{-1}(B) \in \mathcal{F}$，则称 f 为 $(\Omega, \mathcal{F})$ 到 $(E, \mathcal{E})$

的 **可测映射**；若 $\mathcal{F}$ 上有概率测度 $\mathbf{P}$, 则称 f 为 $(\Omega, \mathcal{F}, \mathbf{P})$ 到 $(E, \mathcal{E})$ 的 **随机元**.

如果记 $\{f^{-1}(B): B \in \mathcal{E}\}$ 为 $f^{-1}(\mathcal{E})$, 则 $(\Omega, \mathcal{F})$ 到 $(E, \mathcal{E})$ 的可测映射 f 的条件就可写成如下更简单的符号

$$f^{-1}(\mathcal{E}) \subset \mathcal{F}$$

而将 $f^{-1}(B)$, $f^{-1}(\mathcal{E})$ 分别称为 B、$\mathcal{E}$ 对 f 的 **逆象**，有时也将 $f^{-1}(B)$ 记成 $\{f \in B\}$、将 $\{\omega \in \Omega: |f(\omega)| \leq x\}$ 记成 $\{|f| \leq x\}$，而将 $\mu(\{\omega: f(\omega) \in B\})$ 记成 $\mu(\{f \in B\})$、$\mu(f^{-1}(B))$ 或 $\mu \circ f^{-1}(B)$ 等等.

关于逆象有下述有用性质：

2. 引理 设 f 是由 Ω 到 E 的映射 (不必可测). 则下列公式成立：

(1) $$f^{-1}(E) = \Omega, \qquad f^{-1}(\varnothing) = \varnothing,$$

(2) $$f^{-1}(B^c) = (f^{-1}(B))^c,$$

(3) $$f^{-1}\Big(\bigcup_{\gamma \in \Gamma} B_\gamma\Big) = \bigcup_{\gamma \in \Gamma} f^{-1}(B_\gamma).$$

其中 Γ 为任一指标集 (不必可数), 因而有

(4) $$f^{-1}\Big(\bigcap_{\gamma \in \Gamma} B_\gamma\Big) = \bigcap_{\gamma \in \Gamma} f^{-1}(B_\gamma),$$

(5) $$f^{-1}(B_1 \setminus B_2) = f^{-1}(B_1) \setminus f^{-1}(B_2), \ B_i \subset E, i = 1, 2.$$

证明容易，留给读者.

3. 引理 设 f 是由 Ω 到 E 的映射.

(a) 若 $\mathcal{E}$ 是 E 的一个 σ–代数，则 $f^{-1}(\mathcal{E})$ 是 Ω 的 σ–代数.

(b) 若 $\mathcal{C}$ 是 E 的任一非空的子集类，则

(1) $$f^{-1}(\sigma(\mathcal{C})) = \sigma(f^{-1}(\mathcal{C})).$$

证明 (a) 为引理 2 的直接推论，今往证 (b).

由 (a) 知 $f^{-1}(\sigma(\mathcal{C}))$ 为 Ω 的一个 σ–代数且易见它包含 $f^{-1}(\mathcal{C})$, 故

(2) $$f^{-1}(\sigma(\mathcal{C})) \supset \sigma(f^{-1}(\mathcal{C})).$$

今往证：

(3) $$f^{-1}(\sigma(\mathcal{C})) \subset \sigma(f^{-1}(\mathcal{C})).$$

为此令：

(4) $$\mathcal{A} := \{A \subset E: \ f^{-1}(A) \in \sigma(f^{-1}(\mathcal{C}))\}.$$

显然 $\mathcal{C} \subset \mathcal{A}$, 且 $f^{-1}(\mathcal{A}) \subset \sigma(f^{-1}(\mathcal{C}))$. 于是欲证 (3), 只需证 $\sigma(\mathcal{C}) \subset \mathcal{A}$, 因此只需证 $\mathcal{A}$ 是 E 的一个 σ– 代数即可.

由 (2.1) 知 $f^{-1}(E) = \Omega \in \sigma(f^{-1}(\mathcal{C}))$, 故 $E \in \mathcal{A}$. 若 $A \in \mathcal{A}$, 则由 (4) 知 $f^{-1}(A) \in \sigma(f^{-1}(\mathcal{C}))$, 再由 (2.1) 知 $A^c \in \mathcal{A}$, 其次用同法应用 (2.3) 及 (4) 即知 $\mathcal{A}$ 对可数并封闭，故 $\mathcal{A}$ 为 E 的一个 σ–代数. □

下面的定理，将初等概率论中的 r.v. 的概念与此处的定义 1 联系起来.

4. 定理 X 是可测空间 $(\Omega, \mathcal{F})$ 上的有限实值可测函数 (或随机变量) 当且仅当对一切 $x \in \mathbb{R}$ 有

(1) $$X^{-1}((-\infty, x]) = \{\omega \in \Omega: \ X(\omega) \le x\} \in \mathcal{F}$$

证明 由 (1) 知

(2) $$\sigma(X^{-1}(\{(-\infty, x]: \ x \in \mathbb{R}\})) \subset \mathcal{F}.$$

再由引理 3 知

(3) $$X^{-1}(\sigma(\{(-\infty, x]: \ x \in \mathbb{R}\})) = \sigma(X^{-1}\{(-\infty, x]: \ x \in \mathbb{R}\})$$

而由 Borel σ–代数的定义知 $\mathcal{B} = \sigma(\{(-\infty, x]: \ x \in \mathbb{R}\})$, 故由 (2), (3) 即得 $X^{-1}(\mathcal{B}) \subset \mathcal{F}$, 即 X 是一实可测函数 (或实 r.v.), 即定理的条件是充分的. 条件的必要性显然. □

事实上，这条定理可以大大地推广，在想法上并没有新内容，但却有很多应用. 表达比较复杂一点，现在写在下面，并且给出它对高维 r.v. 的应用.

5. 定理 X 是 $(\Omega,\mathcal{F})$ 到 $(E,\mathcal{E})$ 的可测映射 (或随机元) 的充要条件是存在 $\mathcal{E}$ 的一个子集类 $\mathcal{C}$ 满足：

1) $\sigma(\mathcal{C})=\mathcal{E}$; 2) 对一切 $A\in\mathcal{C}$, $X^{-1}(A)\in\mathcal{F}$.

证明与定理 4 同.

作为定理 5 应用的一个例子，我们讨论与一维 r.v. 结果 (定理 4) 相应的多维情形.

6. 定理 $X=(X_1,\cdots,X_n):\ \Omega\mapsto\mathbb{R}^n$ 是 $(\Omega,\mathcal{F})$ 到 $(\mathbb{R}^n,\mathcal{B}^n)$ 的可测映射的充要条件是对一切 $x=(x_1,\cdots,x_n)\in\mathbb{R}^n$ 有：

$$X^{-1}((-\infty,x])=\{\omega\in\Omega:\ X_k(\omega)\le x_k,k=1,2,\cdots,n\}\in\mathcal{F}$$

证明 在定理 5 中，取 $(E,\mathcal{E})=(\mathbb{R}^n,\mathcal{B}^n)$, $\mathcal{C}=\{(-\infty,x]:\ x\in\mathbb{R}^n\}$ 即可.

定理 6 说明一个 n 维随机变量 (即 n 个随机变量组成的向量) 就是概率空间 $(\Omega,\mathcal{F},\mathbf{P})$ 到 $(\mathbb{R}^n,\mathcal{B}^n)$ 的一个可测映射，即随机元.

下面来讨论 r.v. 的概率分布.

7. 定理 测度空间 $(\Omega,\mathcal{F},\mu)$ 上的每一个 n 维 Borel 可测映射 X 根据下列关系决定一个 n 维 Borel 测度空间 $(\mathbb{R}^n,\mathcal{B}^n,\mu_X)$.

$$(1)\qquad \forall B\in\mathcal{B}^n,\qquad \mu_X(B):=\mu(X^{-1}(B))=\mu(\{X\in B\}).$$

若 μ 是概率，则 μ_X 亦然.

证明 由定理 6 知对一切 $B\in\mathcal{B}^n$, $X^{-1}(B)\in\mathcal{F}$, 所以 $\mu_X(B)$ 有意义，显然 $\mu_X(B)\ge 0$. 设对一切 $B_m\in\mathcal{B}^n$, $n\in\mathbb{N}$, 两两不交，

因而 $X^{-1}(B_m), m \in \mathbb{N}$, 也两两不交. 于是由引理 2,

$$\mu_X\Big(\bigcup_{m=1}^{\infty} B_m\Big) = \mu\Big(X^{-1}\Big(\bigcup_{m=1}^{\infty} B_m\Big)\Big) = \mu\Big(\bigcup_{m=1}^{\infty} X^{-1}(B_m)\Big)$$
$$= \sum_{m=1}^{\infty} \mu(X^{-1}(B_m)) = \sum_{m=1}^{\infty} \mu_X(B_m).$$

当 μ 是概率测度时, 显然 $\mu_X(\mathbb{R}^n) = \mu(X^{-1}(\mathbb{R}^n)) = \mu(\Omega) = 1$. □

8. 注 1) 当 X 是一 n 维随机变量时, $\{X^{-1}(B): B \in \mathcal{B}^n\}$ 是 $\mathcal{F}$ 的一个子 σ–代数. 它只包含与 X 有关的随机事件, 所以通常称为 **由 X 产生的 σ-代数**, 记成 $\sigma(X)$. $\mathbf{P}$ 在 $\sigma(X)$ 上的限制是只与 X 有关的概率规律, 而 $\mathbf{P}_X$ 表示 X 的值的分布的概率规律, 所以称为 X 的 **概率分布测度**, 记作:

$$\mathbf{P}_X = \mathbf{P} \circ X^{-1} \tag{1}$$

由它如下地决定 **X 的概率分布函数**:

$$\begin{aligned} F(x_1, \cdots, x_n) = F(x) &= \mathbf{P}_X((-\infty, x]) = \mathbf{P}(X \le x) \\ &= \mathbf{P}(X_1 \le x_1, \cdots, X_n \le x_n), \\ &\forall x = (x_1, \cdots, x_n) \in \mathbb{R}^n. \end{aligned} \tag{2}$$

以上讨论对复随机变量以至 n 维复随机变量都是适用的, 由上面的讨论知 n 维随机变量 X 唯一决定概率分布测度 $\mathbf{P}_X$, $\mathbf{P}_X$ 唯一决定概率分布函数 F. 反过来, 由第二章推论 2.20 知 F 唯一决定 $\mathbf{P}_X$. 但是 $\mathbf{P}_X$ 并不能唯一决定 X, 即令在空间 $(\mathbb{R}^n, 2^{\mathbb{R}^n})$ 上也不行 (为什么?). 具有相同分布 (测度, 函数) 的随机变量称为 **同分布** 的.

2) 对一般的 n 维 Borel 测度空间的分布函数, 由于有可能 $\mu((-\infty, x]) = \infty$, 所以此时由 $\mathcal{B}^n$ 上的测度 μ 定义分布函数时, 通常采用以下的办法: (还需假定当 B 有界时, $\mu(B) < \infty$. 即 μ

为 Radon 测度!) 为简单起见，只介绍一维情形，设 μ 为 $\mathcal{B}$ 上的测度，任取一参考点 $\theta \in \mathbb{R}$, 定义它的分布函数

$$(3) \qquad F(x) := F_\theta(x) := \begin{cases} \mu((\theta, x]), & x \geq \theta, \\ -\mu((x, \theta]), & x < \theta. \end{cases}$$

于是易见 μ 并不唯一决定分布函数 $F = F_\theta$, 但容易证明凡是满足关系式

$$(4) \qquad \mu((a, b]) = F(b) - F(a), \qquad a, b \in \mathbb{R}^1, \quad a \leq b$$

的函数 F, 相差一个常数，即 μ 与函数类 $\{F + c : c \in \mathbb{R}\}$ 通过关系 (4) 一一对应，其中 F 为一不降右连续函数.

3) 定理 6, 7 不仅对有限维随机变量或可测函数成立，对无穷维的情形 (即随机变量序列，随机过程，随机场的情形) 也成立，不过这些牵涉到无穷维空间及相应 σ–代数的讨论，所以先不介绍了.

4) 定理 7 对一般 $(\Omega, \mathcal{F}, \mu)$ 到距离可测空间 $(E, \mathcal{E})$ 的可测映射也成立，只要将定理 7 中的 $(\mathbb{R}^n, \mathcal{B}^n)$ 换成 $(E, \mathcal{B})$ 即可.

9. 例1 测度空间 $(\Omega, \mathcal{F}, \mu)$, 对一切 $A \in \mathcal{F}$, 定义

$$I_A(\omega) := \begin{cases} 0, & \omega \notin A, \\ 1, & \omega \in A. \end{cases}$$

则 I_A 是 $\mathcal{F}$–可测的函数.

因为对一切 $x \in \mathbb{R}$,

$$\{\omega : I_A(\omega) \leq x\} = \left\{ \begin{array}{ll} \Omega, & x \geq 1; \\ A^c, & 0 \leq x < 1; \\ \varnothing, & x < 0. \end{array} \right\} \in \mathcal{F}.$$

例2 给了分布列

$$\begin{pmatrix} x_1, & x_2, & \cdots \\ p_1, & p_2, & \cdots \end{pmatrix},$$

可以造离散 r.v. 如下：令 $\Omega := \{x_n : n \in \mathbb{N}\}, \mathcal{F} := \{A : A \subset \Omega\}$

$$P(A) := \sum_{x_n \in A} p_n \qquad \forall A \subset \Omega.$$

$$X(x_n) := x_n.$$

则 X 是服从上述分布列的离散随机变量. 因而可以造服从二项分布，Poisson 分布等的 r.v..

例3 给了非负可测函数 $f : \mathbb{R} \longrightarrow \mathbb{R}_+$ 且满足 $\int_{-\infty}^{\infty} f(x)\,dx = 1$ (此处及下面一行的积分是 Lebesgue 积分，将在下一章讨论). 可以造连续型 r.v. 如下：令 $\Omega = \mathbb{R}, \mathcal{F} = \mathcal{B}, \mathbf{P}(A) = \int_A f(x)\,dx, A \in \mathcal{B}$(由下一章的积分理论知，$\mathbf{P}$ 是 $\mathcal{B}$ 上的概率测度)

$$X(x) := x, \qquad x \in \mathbb{R}.$$

则 X 是以 f 为概率分布密度的随机变量. 当

$$f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & x < 0 \text{或} x > 1 \end{cases}$$

时，X 为服从均匀分布 $U[0,1]$ 的 r.v.. 当

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

时，X 为服从正态分布 $N(\mu, \sigma^2)$ 的 r.v..

下面我们讨论可测映射的复合映射仍然可测，以及它的特例多维随机变量的函数仍是随机变量的问题.

10. 定理 设 $(\Omega, \mathcal{F}), (E, \mathcal{E}), (A, \mathcal{A})$ 是可测空间，X 是 $(\Omega, \mathcal{F})$ 到 $(E, \mathcal{E})$ 的可测映射，Y 是 $(E, \mathcal{E})$ 到 $(A, \mathcal{A})$ 的可测映射，则 $Y \circ X$ 为 $(\Omega, \mathcal{F})$ 到 $(A, \mathcal{A})$ 的可测映射，其中 $(Y \circ X)(\omega) := Y(X(\omega))$.

证明 由假设知 $X^{-1}(\mathcal{E}) \subset \mathcal{F}, Y^{-1}(\mathcal{A}) \subset \mathcal{E}$, 所以由 $Y \circ X$ 的定义知

$$(Y \circ X)^{-1}(\mathcal{A}) == X^{-1}(Y^{-1}(\mathcal{A})) \subset X^{-1}(\mathcal{E}) \subset \mathcal{F}. \quad \square$$

11. 推论 1) 若 X 为 n 维有限值 r.v., f 是 $\mathbb{R}^n$ 到 $\mathbb{R}$ 上的连续函数 (可以是复值的), 则 $f(X)$ 是 r.v..

特别地, X, Y 是有限值 r.v., 则 $X^r(r \in \mathbb{N})$, $|X|^r(r \in \mathbb{R}_+)$, $e^{-\lambda X}(\lambda \in \mathbb{R})$, $e^{itX}(t \in \mathbb{R})$, $X \vee Y$, $X \wedge Y$, $X+Y$, $X-Y$, XY, $X/Y(Y(\omega) \neq 0$, 对一切 $\omega \in \Omega)$ 都是 r.v., 其中 $x \vee y := \max(x, y)$, $x \wedge y := \min(x, y)$.

2) 若 X 为 n 维 r.v., f 是 $\overline{\mathbb{R}}^n$ 到 $\overline{\mathbb{R}}$ 上的连续函数, 则 $f(X)$ 是 r.v..

特别地, X, Y 是 r.v., 则 $X^r(r \in \mathbb{N})$, $|X|^r(r \in \mathbb{R}_+)$, $e^{-\lambda X}(\lambda \in \mathbb{R})$, $X \vee Y$, $X \wedge Y$, XY 都是 r.v., 其中 $x \vee y := \max(x, y)$, $x \wedge y := \min(x, y)$. 若 $X+Y, X-Y, X/Y$ 有意义, 也都是可测函数.

证明 因为若 $f: \mathbb{R}^n \to \mathbb{R}$ 连续, 则由引理 2.5.2(2) 知对一切 $G \subset \mathbb{R}$ 开集, $f^{-1}(G)$ 为 $\mathbb{R}^n$ 中的开集, 因而属于 $\mathcal{B}^n$, 记 $\mathbb{R}$ 中的一切开子集类为 $\mathcal{O}$, 则 $f^{-1}(\mathcal{O}) \subset \mathcal{B}^n$, 于是由 $\sigma(\mathcal{O}) = \mathcal{B}$ 及定理 5 知 $f^{-1}(\mathcal{B}) \subset \mathcal{B}^n$, 即 f 为 n 元 Borel 函数. 由定理 10 知推论的结论 (1) 成立. 结论 (2) 的证明类似. □

定理 10 及推论 11 从理论上说明了可测映射的复合映射及特例多维随机变量的可测 (连续) 函数仍然是随机变量, 这一点不用测度论是说不清楚的.

下面讨论可数个 r.v. 的情形, 主要是关于函数序列的极限的可测性问题. 首先我们复习有关函数的极限的概念.

12. 定义 设 $\{f_n: n \in \mathbb{N}\}$ 是定义在 Ω 上的广义实值可测函数列, 对一切 $\omega \in \Omega$, 定义

$$(1) \qquad \left(\sup_n f_n\right)(\omega) := \sup_n f_n(\omega); \qquad \left(\inf_n f_n\right)(\omega) = \inf_n f_n(\omega);$$

$$(2) \qquad \left(\overline{\lim_{n\to\infty}} f_n\right)(\omega) = \inf_n \left(\sup_{k \geq n} f_k(\omega)\right);$$

$$(3)\qquad \left(\varliminf_{n\to\infty} f_n\right)(\omega) := \sup_n \left(\inf_{k\geq n} f_k(\omega)\right).$$

则 $\sup\limits_n f_n$, $\inf\limits_n f_n$, $\varlimsup\limits_{n\to\infty} f_n$, $\varliminf\limits_{n\to\infty} f_n$ 分别称为序列 $\{f_n:\ n\in\mathbb{N}\}$ 的**上确界、下确界、上极限、下极限**.

习知 $\lim\limits_{n\to\infty} f_n(\omega)$ 存在 (包括为 $\pm\infty$) 当且仅当 $\varliminf\limits_{n\to\infty} f_n(\omega) = \varlimsup\limits_{n\to\infty} f_n(\omega)$, 因此定义 f_n 的极限函数 $\lim_{n\to\infty} f_n$ 为

$$(4)\qquad \Delta := \left\{\omega:\ \varliminf_{n\to\infty} f_n(\omega) = \varlimsup_{n\to\infty} f_n(\omega)\right\}$$

上的广义实值函数，其定义为：对一切 $\omega\in\Delta$,

$$(5)\qquad \left(\lim_{n\to\infty} f_n\right)(\omega) := \lim_{n\to\infty} f_n(\omega).$$

13. 定理 若 $\{f_n:\ n\in\mathbb{N}\}$ 是 $(\Omega,\mathcal{F})$ 上的 (广义) 实可测函数序列，则 $\inf_n f_n$, $\sup_n f_n$, $\varliminf_{n\to\infty} f_n$, $\varlimsup_{n\to\infty} f_n$ 都是 $(\Omega,\mathcal{F})$ 上的 (广义) 实可测函数，且由 (12.4) 定义的 $\Delta\in\mathcal{F}$, 而 $\lim_{n\to\infty} f_n$ 是 Δ 上的 (广义) 实可测函数.

注：由最后结论看出，定义 1 中一般可测函数概念的必要性.

证明 显然对一切 $x\in\mathbb{R}$,

$$\left\{\sup_n f_n \leq x\right\} = \bigcap_{n\in\mathbb{N}} \{f_n \leq x\},$$

$$\left\{\inf_n f_n \geq x\right\} = \bigcup_{n\in\mathbb{N}} \{f_n \geq x\}.$$

故 $\sup_n f_n$, $\inf_n f_n$ 都是 $\mathcal{F}$-可测的，因而 $\varliminf_{n\to\infty} f_n$, $\varlimsup_{n\to\infty} f_n$ 也是.

为了证明 $\Delta\in\mathcal{F}$, 我们注意对任何两个实 $\mathcal{F}$-可测函数 f, g 来说，

$$\{\omega:\ f(\omega)\neq g(\omega)\} = \{\omega:\ f(\omega) < g(\omega)\}\cup\{\omega:\ f(\omega) > g(\omega)\}\in\mathcal{F}.$$

设 Q 表示有理数集，因为

$$\{\omega:\ f(\omega)<g(\omega)\}=\bigcup_{r\in Q}(\{\omega:\ f(\omega)<r\}\cap\{\omega:\ r<g(\omega)\})\in\mathcal{F},$$

$$\{\omega:\ f(\omega)>g(\omega)\}=\bigcup_{r\in Q}(\{\omega:\ f(\omega)>r\}\cap\{\omega:\ g(\omega)<r\})\in\mathcal{F},$$

所以前述诸结论成立. □

这条定理说明可测函数在进行极限运算时较连续函数方便.

14. **注** 应用本节的结论容易给出随机变量的可测 (或连续) 函数的分布的一般公式，如果原来随机变量的分布已知的话.

设 X 是一 n 维实随机变量， μ_X 是它的分布， $f:\mathbb{R}^n\mapsto\mathbb{R}^m$ 是 Borel 可测函数，则 $f(X)$ 的概率分布测度及分布函数可分别表示如下：

(1) $\mu_{f(X)}=\mu_X\circ f^{-1}$，即 $\mu_{f(X)}(B)=\mu_X(f^{-1}(B))\quad\forall B\in\mathcal{B}^m$.

(2) $$F_{f(X)}(x)=\mu_X(f^{-1}((-\infty,x])),\quad\forall x\in\mathbb{R}^m.$$

例 设一维随机变量 X 的分布函数为 F, 则 $\sin X$ 的分布函数的值

$$\begin{aligned}F_{\sin X}(x)&=\mu_X(\sin^{-1}([-1,x]))\\&=\sum_{n=-\infty}^{\infty}\mu_X([(2n-1)\pi-\sin^{-1}x,2n\pi+\sin^{-1}x])\\&=\sum_{n=-\infty}^{\infty}[F(2n\pi+\sin^{-1}x)-F((2n-1)\pi-\sin^{-1}x-0)],\\&\quad\forall x\in[-1,1];\end{aligned}$$

此外， $F_{\sin X}(x)=0$, 对一切 $x<-1$; $F_{\sin X}(x)=1$, 对一切 $x>1$.

习题

1. 试证示性函数有下列性质：

1) $I_{A\cap B}=I_A\cdot I_B$, $I_{A\cup B}=I_A+I_B-I_{A\cap B}$, $I_{A^c}=1-I_A$, $I_{A\Delta B}=|I_A-I_B|$;

2) $I_{\cup_n A_n}=\sup_n I_{A_n}$, $I_{\cap_n A_n}=\inf_n I_{A_n}$,

$I_{\overline{\lim}_{n\to\infty}A_n}=\overline{\lim}_{n\to\infty} I_{A_n}$,

$I_{\underline{\lim}_{n\to\infty}A_n}=\underline{\lim}_{n\to\infty} I_{A_n}$.

2. 设 f, g 是 $(\Omega,\mathcal{F})$ 上的实 (或复) 可测函数，问下列函数是否 $\mathcal{F}$可测函数？并说明理由.

1) $f_1(\omega):=\begin{cases} f(\omega), & \omega\in\{|f|<\infty\},\\ 0, & \omega\in\{|f|=\infty\}\end{cases}$；即 $f_1=f\cdot I_{\{|f|<\infty\}}$;

2) $f_2:=f\cdot I_{\{\omega\}^c}+(f+1)I_{\{\omega\}}$, ω 为 Ω 的给定元；

3) $f_3:=f\cdot I_A+g\cdot I_{A^c}$, $A\in\mathcal{F}$.

3. f 是实 $(\Omega,\mathcal{F})$ 可测函数，则 $|f|$ 是 $\mathcal{F}$ 可测的，逆命题是否成立？

4. 设 $\mathcal{F}:=\{\emptyset, A, A^c, \Omega\}$, $A\subset\Omega$ 给定. 试写出 $(\Omega,\mathcal{F})$ 上的全部 $\mathcal{F}$–可测函数.

5. 如果两个随机变量几乎处处相等，则它们具有相同的概率分布测度.

6. 对于给定的 $(\mathbb{R},\mathcal{B})$ 上的任何概率测度 μ, 定义一个概率分布测度为 μ 的随机变量.

7. 设 θ 为 $[0,1]$ 中的均匀分布，对于每个分布函数 F, 定义 $G(y)=\sup\{x:\ F(x)\le y\}$, 则 $G(\theta)$ 具有分布函数 F.

8. 设 X 具有连续分布函数 F, 则 $F(X)$ 具有 $[0,1]$ 上的均匀分布，如果 $F(x)$ 不连续，情况又怎样？

9. 如果 f 为 Borel 可测， X 与 Y 同分布，则 $f(X)$ 与 $f(Y)$ 也同分布.

10. 设 $\sigma(X)$ 是由随机变量 X 所产生的 $\sigma-$ 域，则 $\Lambda \in \sigma(X)$ 的充要条件是对于某个 $B \in \mathcal{B}$, $\Lambda = X^{-1}(B)$. 问此 B 是否唯一？能否存在一个集 $A \notin \mathcal{B}$, 使得 $\Lambda = X^{-1}(A)$？

11. 将习题 10 中的论断推广到有限个随机变量的情况. （推广到任意个随机变量的情况也是可能的）

12. 设 X 是 n 维实随机变量，F_X, μ_X 分别是 X 的分布函数和概率分布测度. 试用 F_X, μ_X 表示 $f(X)$ 的分布函数或概率分布测度，其中 $f: \mathbb{R}^n \to \mathbb{R}$ 是下列 Borel 可测函数：

1) 当 $n=1$ 时，$f(x) := x^2$, $x \in \mathbb{R}$;

2) 当 $n=1$ 时，$f(x) := (x+4)(x-1)(x-3)$, $x \in \mathbb{R}$;

3) 当 $n=1$ 时，$f(x) := \cos kx$, $x \in \mathbb{R}$, k 为常数；

4) $f(x) := \sqrt{\sum_{k=1}^n x_k^2}$, $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$;

5) $f(x) := \sum_{k=1}^n x_k$, $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$;

6) $f(x) := \max_{1 \le k \le n} x_k$, $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$;

7) $f(x) := \min_{1 \le k \le n} x_k$, $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$.

§2. 可测函数的构造性质

1. 定义 给定 $(\Omega, \mathcal{F})$.

1) 若 $A_k \in \mathcal{F}$, $k = 1, 2, \cdots, n$, 两两不交且 $\bigcup_{k=1}^n A_k = \Omega$, a_k, $k = 1, 2, \cdots, n$ 为实数或 $+\infty$（或复数），称函数

$$f := \sum_{k=1}^n a_k I_{A_k} \qquad (\text{即 } f(\omega) := \sum_{k=1}^n a_k I_{A_k}(\omega), \quad \omega \in \Omega)$$

为 $\mathcal{F}$-**简单函数**.

2) 若 $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, 两两不交且 $\bigcup_{n \in \mathbb{N}} A_n = \Omega$, a_n, $n \in \mathbb{N}$ 为广义实数（或复数），则称函数

$$f := \sum_{n=1}^\infty a_n I_{A_n}$$

为 $\mathcal{F}$–**初等函数**.

2. 引理 设 $(\Omega,\mathcal{F})$ 为可测空间, 则 $\mathcal{F}$–简单函数, $\mathcal{F}$– 初等函数都是 $\mathcal{F}$–可测函数.

证明 由 1.9 例 1 知 I_A, $A\in\mathcal{F}$ 是 $\mathcal{F}$–可测的, 因而 $\mathcal{F}$– 简单函数的 $\mathcal{F}$–可测性由定理 1.11 立得. 而 $\mathcal{F}$– 初等函数的可测性则由定理 1.13 可得 (用到 $\sum_{n=1}^{\infty} a_n I_{A_n} = \lim_{n\to\infty}\sum_{k=1}^{n} a_k I_{A_k}$).

下面给出另一种证明.

设 $A_n\in\mathcal{F}$, $n\in\mathbb{N}$, 两两不交, a_n 为广义实数 (或复数). $f=\sum_{n=1}^{\infty} a_n I_{A_n}$, 则 $f(\omega)=a_n$, 当 $\omega\in A_n$. 于是对一切 $B\in\mathcal{B}$, 有

$$f^{-1}(B)=\{\omega\in\Omega:\ f(\omega)\in B\}=\bigcup_{a_n\in B}\{\omega\in\Omega:\ f(\omega)=a_n\}$$
$$=\bigcup_{n:\ a_n\in B} A_n\in\mathcal{F}.$$

因此 f 为 $\mathcal{F}$–可测函数. □

例 $(\Omega,\mathcal{F},\mathbf{P})$ 是概率空间, $A_n\in\mathcal{F}$, $n\in\mathbb{N}$, 两两不交, $\bigcup_{n=1}^{\infty} A_n=\Omega$. 令

$$X(\omega)=x_n,\qquad 当\ \omega\in A_n,\quad 即 X:=\sum_{n=1}^{\infty} x_n I_{A_n},$$

则 X 为一离散 r.v., 取值 $\{x_n:\ n\in\mathbb{N}\}$, 且分布列为

$$p_n:=\mathbf{P}(\{X=x_n\})=\mathbf{P}(A_n).$$

3. 引理 设 $(\Omega,\mathcal{F})$ 为可测空间, f,g 为任意两个广义实 $\mathcal{F}$–简单函数 ($\mathcal{F}$– 初等函数), 则

1) $f+ig$ 为 $\mathcal{F}$– 简单函数 (相应地: $\mathcal{F}$– 初等函数);

2) 当 $f+g$ 有意义时, $f+g$ 为 $\mathcal{F}$ 简单函数 (相应地: $\mathcal{F}$– 初等函数).

证明 设 $f=\sum_{n=1}^{\infty}a_nI_{A_n}$, $g=\sum_{m=1}^{\infty}b_mI_{B_m}$, 则

$$f+ig=\sum_{n,m=1}^{\infty}(a_n+ib_m)I_{A_nB_m},$$

$$f+g=\sum_{n,m=1}^{\infty}(a_n+b_m)I_{A_nB_m}. \quad \square$$

现在我们研究可测函数的一种构造性质 — 用简单函数来逼近.

4. **定理** 设 $(\Omega,\mathscr{F})$ 是可测空间, 则

1) 任一 $\mathscr{F}$–可测函数是 $\mathscr{F}$–简单函数序列的极限;

2) 任一 $\mathscr{F}$–可测函数是 $\mathscr{F}$–初等函数序列的一致极限;

3) 任一有界 $\mathscr{F}$–可测函数是 $\mathscr{F}$–简单函数序列的一致极限;

4) 任一非负 $\mathscr{F}$–可测函数是非负不降 $\mathscr{F}$–简单函数 ($\mathscr{F}$–初等函数) 序列的极限 (相应地: 一致极限).

证明 由引理 3 的 1) 知只需证实值函数的情形. 设 f 为实 $\mathscr{F}$–可测函数.

1) 对一切 $n\in\mathbb{N}$, 令

$$f_n:=\sum_{k=-n\cdot 2^n}^{n\cdot 2^n-1}\frac{k}{2^n}I_{\{\frac{k}{2^n}\le f<\frac{k+1}{2^n}\}}+nI_{\{f\ge n\}}+(-n)I_{\{f<-n\}}, \tag{1}$$

则显然 f_n 是 $\mathscr{F}$–简单函数, 且

$$|f_n(\omega)-f(\omega)|<\frac{1}{2^n}, \qquad \text{当 } -n\le f(\omega)<n \text{ 时}$$

$$f_n(\omega)=n, \qquad \text{当 } f(\omega)=+\infty \text{ 时;}$$

$$f_n(\omega)=-n, \qquad \text{当 } f(\omega)=-\infty \text{ 时.}$$

由此即知对一切 $\omega\in\Omega$, 有

$$\lim_{n\to\infty}f_n(\omega)=f(\omega).$$

即 1) 获证.

2) 为证 2), 对一切 $n\in\mathbb{N}$, 令

$$(2)\quad g_n := \sum_{k=-\infty}^{\infty}\frac{k}{2^k}I_{\{\frac{k}{2^n}\le f<\frac{k+1}{2^n}\}}+\infty\cdot I_{\{f=+\infty\}}+(-\infty)I_{\{f=-\infty\}},$$

则 g_n 为 $\mathcal{F}$–初等函数, 且

$$|g_n(\omega)-f(\omega)|<\frac{1}{2^n},\qquad 当 f(\omega)\in\mathbb{R} 时;$$

$$g_n(\omega)=f(\omega),\qquad 当 f(\omega)=\pm\infty 时.$$

由此即知, 当 $n\to\infty$ 时, g_n 一致地趋于 f(实际上, 对一切 $\omega\in\Omega$, $f(\omega)-\frac{1}{2^n}<g_n(\omega)<f(\omega)+\frac{1}{2^n}$).

3) 若存在 $M>0$, 使 $|f(\omega)|<M$, 对一切 $\omega\in\Omega$ 成立时, 则由 $n>M$ 时, 由 1) 定义的 f_n 没有最后两项, 因而当 $n>M$ 时,

$$|f_n(\omega)-f(\omega)|<\frac{1}{2^n},\qquad \forall\omega\in\Omega.$$

故当 $n\to\infty$, f_n 一致地趋于 f.

4) 当 f 为非负 $\mathcal{F}$–可测时, 则对一切 $n\in\mathbb{N}$, 令

$$f_n := \sum_{k=0}^{n\cdot 2^n-1}\frac{k}{2^n}I_{\{\frac{k}{2^n}\le f<\frac{k+1}{2^n}\}}+nI_{\{f\ge n\}},$$

$$g_n := \sum_{k=0}^{\infty}\frac{k}{2^n}I_{\{\frac{k}{2^n}\le f<\frac{k+1}{2^n}\}},$$

则 $\{f_n:\ n\in\mathbb{N}\}(\{g_n:\ n\in\mathbb{N}\})$ 为非负不降 $\mathcal{F}$–简单函数 ($\mathcal{F}$–初等函数) 序列, 且

$$\lim_{n\to\infty}f_n(\omega)=f(\omega),\qquad \forall\omega\in\Omega 成立$$

$$\lim_{n\to\infty}g_n(\omega)=g(\omega),\qquad \forall\omega\in\Omega 一致成立.\quad\square$$

5. 定义 设 $f:\Omega\to\overline{\mathbb{R}}$. 分别称 $f^+=f\vee 0$, $f^-=(-f)\vee 0=-(f\wedge 0)$ 为 f 的 **正部** 与 **负部**.

显然,

$$f = f^+ - f^-, \tag{1}$$

$$|f| = f^+ + f^-, \tag{2}$$

而当 f 为有限值函数时,

$$f^+ = \frac{|f|+f}{2}, \qquad f^- = \frac{|f|-f}{2}. \tag{3}$$

6. 定理 实可测函数的正部与负部都是可测函数，因而实可测函数可以表成二非负可测函数之差.

证明 定理的前一部分即推论 1.11 的一部分. 而后一部分由 (1) 即得. □

下面将介绍的可测函数的性质，也可以认为是可测函数的另一种构造性质. 为此我们先介绍一个函数形式的单调类定理，它是测度论的一个很重要的工具.

7. 定义 设 $\mathcal{L}$ 是定义在 Ω 上的广义实值函数类，满足条件: $f \in \mathcal{L} \implies f^+, f^- \in \mathcal{L}$.

函数族 L 称为 $\mathcal{L}$–系. 如果满足:

I) $1 \in L$;

II) L 中有限个函数的线性组合 (如果有意义) 属于 L;

III) 若 $f_n \in L$, $n \in \mathbb{N}$, $0 \leq f_n \uparrow f$, f 有界或 $f \in \mathcal{L}$, 则 $f \in L$.

8. 定理 (函数形式的单调类定理) 若 $\mathcal{L}$–系 L 包含一 π–系 $\mathcal{C}$ 中任一集的示性函数，则 L 包含一切属于 $\mathcal{L}$ 的 $\sigma(\mathcal{C})$–可测函数.

证明 令

$$\Lambda := \{A \subset \Omega : \ I_A \in L\}.$$

则由 I)—III) 知 $\Omega \in \Lambda$, Λ 对真差封闭且对不降集列的并封闭，因而 Λ 为一 λ– 系，由定理假设知 $\Lambda \supset \mathcal{C}$. $\mathcal{C}$ 为 π–系，所以由集合

形式的单调类定理知 $\mathcal{A}\supset\sigma(\mathcal{C})$, 因而

$$\{I_A:\ A\in\sigma(\mathcal{C})\}\subset L.$$

再由 II) 知任何 $\sigma(\mathcal{C})$– 简单函数属于 L.

设 f 为 $\mathcal{L}$ 中非负 $\sigma(\mathcal{C})$–可测函数, 则由定理 4(4) 知存在非负不降 $\sigma(\mathcal{C})$–简单函数列 $f_n\uparrow f$. 于是由 III) 知 $f\in L$, 若 $f\in\mathcal{L}$ 且 $\sigma(\mathcal{C})$– 可测, 则由 $\mathcal{L}$ 的定义知 $f^+,f^-\in\mathcal{L}$, 且非负 $\sigma(\mathcal{C})$–可测, 于是由前知 $f^+,f^-\in L$, 而 $f=f^+-f^-$ 有意义, 故由 II) $f\in L$, 即定理获证. □

定理的典型用法如下: 要想证明某一函数族 F 具有某种性质 A_0, 为此目的引入一个函数族 $\mathcal{L}$ (满足定义 7 的条件), 使:

$$L:=\{f:\ \text{函数}f\text{具性质}A_0\}$$

为一 $\mathcal{L}$–系, 再引入一个 π–系 $\mathcal{C}$, 使 $\mathcal{L}$ 中的 $\sigma(\mathcal{C})$–可测函数类包含 F, 如果做到这一切, 那么根据定理 8, 只要证明对一切 $A\in\mathcal{C}$, $I_A\in L$ 就够了. 这个方法称为 $\mathcal{L}$**–系方法**.

下面我们应用 $\mathcal{L}$–系方法来证明几个有用的定理.

9. 定理 设 Ω 为一集合, $(E,\mathcal{E})$ 为可测空间, $f:\ \Omega\mapsto E$, 令 $\sigma(f)=f^{-1}(\mathcal{E})$ (它是 Ω 的一个 σ–代数). 则 φ 为从 Ω 到 $\overline{\mathbb{R}}^1$ 的 $\sigma(f)$–可测函数的充分必要条件是有 $(E,\mathcal{E})$ 上的广义实值可测函数 g 存在, 使得 $\varphi=g\circ f$, 并且若 φ–有限 (有界), 则可取 g 有限 (相应地, 有界).

证明 如果 $\varphi=g\circ f$, g 是 $(E,\mathcal{E})$ 到 $(\overline{\mathbb{R}}^1,\overline{\mathcal{B}}^1)$ 的可测函数, 由定理 1.10 知 $g\circ f$ 是 $(\Omega,\sigma(f))$ 到 $(\overline{\mathbb{R}}^1,\overline{\mathcal{B}}^1)$ 的可测映射, 即 φ 为 $\sigma(f)$–可测函数. 充分性获证.

往证条件的必要性: 令

$$\mathcal{L}:=\{h:\ \Omega\mapsto\overline{\mathbb{R}}^1\},$$

$$L:=\{g\circ f:\ g\in\mathcal{E}\}.$$

(为书写简单起见, 今后用 $g \in \mathcal{E}$ 表示 g 为 $\mathcal{E}$–可测函数.) 首先证明: L 是 $\mathcal{L}$–系. 由于 $\mathbf{1} = I_E \circ f$, $I_E \in \mathcal{E}$, 因而 $\mathbf{1} \in L$. 再有对一切 $\alpha \in \mathbb{R}$, $\alpha(g \circ f) = (\alpha g) \circ f$, 知 L 对数乘封闭, 若 $g_1 \circ f$, $g_2 \circ f \in L$, 且 $g_1 \circ f + g_2 \circ f$ 有意义, 则对一切 $\omega \in \Omega$,

$$f(\omega) \notin A := \{x \in E:\ g_1(x) = -g_2(x) = \pm\infty\}.$$

(否则 $(g_1 \circ f + g_2 \circ f)(\omega) = g_1(f(\omega)) + g_2(f(\omega))$ 无意义!) 于是令

$$g := g_1 + g_2 I_{A^c},$$

则 $g \in \mathcal{E}$ 且 $g_1 \circ f + g_2 \circ f = g \circ f$. 因而 $g_1 \circ f + g_2 \circ f \in L$. 故 II) 成立. 最后, 若 $\varphi_n \in L$, $0 \le \varphi_n \uparrow \varphi$, 则存在 $g_n \in \mathcal{E}$, 使 $\varphi_n = g_n \circ f$, 令 $g := \sup_n g_n$, 则 $\varphi = g \circ f$ 且 $g \in \mathcal{E}$. 因而 $\varphi \in L$, III) 成立, 故 L 为一 $\mathcal{L}$–系.

于是为证条件的必要性, 由单调类定理只需证: L 包含 $\sigma(f)$ 的元的示性函数. 若 $C \in \sigma(f)$, 则存在 $B \in \mathcal{E}$, 使 $C = f^{-1}(B)$, 于是易见 $I_B \in \mathcal{E}$,

$$I_C = I_B \circ f \in L.$$

故必要性获证.

若 φ 有限, 且 $\varphi = g \circ f$, 则 $f(\Omega) \subset \{|g| < \infty\}$, 令 $\bar{g} := g \cdot I_{\{|g|<\infty\}}$, 于是 $\bar{g}$ 有限, $\bar{g} \in \mathcal{E}$ 且 $\varphi = \bar{g} \circ f$.

若存在 $M < \infty$ 使 $|\varphi| < M$, 且 $\varphi = g \circ f$, 则令 $\tilde{g} := g \circ I_{\{|g|<M\}}$, 则 $\tilde{g}$ 有界, $\tilde{g} \in \mathcal{E}$, 且 $\varphi = \tilde{g} \circ f$.

至此定理证毕. □

10. 推论 $f = (f_1, \cdots, f_n)$ 是 Ω 上的 n 维实可测函数, 则 g 为 $f^{-1}(\overline{\mathcal{B}}^n)$–可测的充要条件是存在 $(\overline{\mathbb{R}}^n, \overline{\mathcal{B}}^n)$ 上的可测函数 G, 使对一切 $\omega \in \Omega$, 有

$$g(\omega) = G(f_1(\omega), \cdots, f_n(\omega)),$$

还有与推论 10 相应的关于随机过程的结论，我们就不再在此叙述了.

11. 定理 设 $\mathcal{L}$ 是定义在 $\overline{\mathbb{R}}^n$ 上的实函数类， L 是 $\overline{\mathbb{R}}^n$ 上包含一切有界连续函数的 $\mathcal{L}$–系，则 L 包含 $\overline{\mathbb{R}}^n$ 上一切 Borel 可测函数.

证明 令

$$\mathcal{C} := \{A:\ A\text{是}\overline{\mathbb{R}}^n\text{中有限或无限开区间}\},$$

则 $\mathcal{C}$ 是 π–系，且 $\sigma(\mathcal{C}) = \overline{\mathcal{B}}^n$. 故若能证明：对一切 $A \in \mathcal{C}$, $I_A \in L$, 则由定理 8 知此定理获证.

设 $A \in \mathcal{C}$ 为有限区间，则可设

$$A = \{(x_1, \cdots, x_n):\ a_k < x_k < b_k,\ k = 1, 2, \cdots, n\},$$

其中 $a_k, b_k \in \mathbb{R}^1$, $a_k < b_k$, $k = 1, 2, \cdots, n$.

$$f^{(\ell)}(x) = \prod_{k=1}^{n} f_k^{(\ell)}(x_k), \qquad x = (x_1, \cdots, x_n) \in \overline{\mathbb{R}}^n, \quad \ell \in \mathbb{N},$$

其中

$$f_k^{(\ell)}(x_k) = \begin{cases} 0, & x_k \le a_k\text{或}x_k \ge b_k, \\ 1, & a_k + \ell^{-1} \le x_k \le b_k - \ell^{-1}, \\ \ell(x_k - a_k), & a_k < x_k < a_k + \ell^{-1}, \\ \ell(b - x_k), & b_k - \ell^{-1} < x_k < b_k. \end{cases}$$

则易见 $f^{(\ell)}(x)$ 是 $\overline{\mathbb{R}}^n$ 上的有界连续函数，且当 $\ell \uparrow \infty$ 时， $f^{(\ell)}(x) \uparrow I_A(x)$, $x \in \overline{\mathbb{R}}^n$. 故由定义 7(III) 知 $I_A \in L$. 对于 A 为 $\overline{\mathbb{R}}^n$ 中的无限开区间的情况，也可类似地证明 $I_A \in L$. 故定理获证.

习题

1. 设 $\mathcal{C}$ 为 Ω 的一个 π–系， $\mathcal{H}$ 为 Ω 上的函数类，且满足下列条件：

1) $\mathbf{1} \in \mathcal{H}$;

2) $\mathcal{H}$ 对非负线性组合封闭，且若 $f, g \in \mathcal{H}$, 有界， $f \geq g$, 则 $f - g \in \mathcal{H}$;

3) 若 $f_n \in \mathcal{H}$, $n \in \mathbb{N}$, 且 $0 \leq f_n \uparrow f$, 则 $f \in \mathcal{H}$;

4) $\mathcal{H} \supset \{I_A : A \in \mathcal{C}\}$.

则 $\mathcal{H}$ 包含一切非负 $\sigma(\mathcal{C})$–可测函数.

2. 设 $f: \mathbb{R} \mapsto \mathbb{R}$ 在 $\mathbb{R}$ 的每一点可求导，试证其导函数 Borel 可测.

3. Cantor 集 $P_0 = [0,1] \setminus G_0$,

$$G_0 := \bigcup_{n=1}^{\infty} \bigcup_{\substack{a_k=0,2 \\ k=1,2,\cdots n-1}} \left(\frac{1}{3^n}\left[\sum_{k=1}^{n-1} a_{n-k}3^k + 1\right], \frac{1}{3^n}\left[\sum_{k=1}^{n-1} a_{n-k}3^k + 2\right] \right).$$

今定义函数 $f: [0,1] \mapsto [0,1]$ 如下：

当 $x \in \left(\frac{1}{3^n}[\sum_{k=1}^{n-1} a_{n-k}3^k + 1], \frac{1}{3^n}[\sum_{k=1}^{n-1} a_{n-k}3^k + 2]\right)$ 时,

$$f(x) := \frac{1}{2^n}\left(\sum_{k=1}^{n-1} \frac{a_{n-k}}{2} \cdot 2^k + 1\right).$$

于是 f 在 G_0 上有定义，且在 G_0 上不降，其次，令 $f(0) = 0$, 而对一切 $x \in P_0 \setminus \{0\}$, 定义

$$f(x) := \sup\{f(t) : t \in G_0, \quad t < x\}.$$

试证： f 为 $[0,1]$ 上的不降连续函数，因而 f 为 $\mathcal{B}[0,1]$–可测.

4. 设 (E,d) 是一距离空间，试证：

1) 对一切 $G \subset E$ 为开集，令 $f_n := \dfrac{nd(x,G^c)}{1+nd(x,G^c)}$, $x \in E$, 则 $0 \leq f_n \uparrow I_G$;

2) 设 $\mathcal{L}$ 为 (E,d) 上的全体实值函数类， L 为 $\mathcal{L}$–系，且 $L \supset C_b(E,\mathbb{R})$, (即 E 上的有界实连续函数类). , 则 L 包含 $\mathcal{B}(E)$–可测函数类.

第五章 积分与数学期望

§1. 积分的定义

1. 在概率论中，若 X 为概率空间 $(\Omega,\mathcal{F},\mathbf{P})$ 上的离散 r.v.，设 X 取有限或可数个值 $\{x_n\}$，且 $\mathbf{P}(\{X=x_n\})=p_n,\ n\in I,I=\mathbb{N}$ 或 $\{1,\cdots,N\}$，则由概率论知 X 的数学期望 (当下述级数满足一定条件时) 为

$$\mathbf{E}X:=\sum_{n\in I}x_np_n. \tag{1}$$

上述随机变量可以写成 $X=\sum_n x_nI_{\{X=x_n\}}=\sum_n x_nI_{A_n}$，其中 $A_n:=\{X=x_n\}$，则 (1) 即为

$$\mathbf{E}X=\sum_n x_n\mathbf{P}(X=x_n)=\sum_n x_n\mathbf{P}(A_n). \tag{2}$$

实际上，当 X 取有限个 (可数个) 值时，(2) 的右边就是 $\mathcal{F}-$ 简单函数 ($\mathcal{F}-$ 初等函数) 关于 $\mathbf{P}$ 的积分，因此给出以下

2. 定义 设 $(\Omega,\mathcal{F},\mu)$ 为一测度空间，

1) 若 f 为非负 $\mathcal{F}-$ 简单函数，设

$$f=\sum_{k=1}^m x_kI_{A_k},x_k\geqslant 0,A_k\in\mathcal{F},A_1,\cdots,A_m,\text{两两不交},\bigcup_{k=1}^m A_k=\Omega,$$

则定义 f **在 Ω 上对 μ 的积分**

$$\int_\Omega f\,d\mu:=\sum_{k=1}^m x_k\mu(A_k).$$

2) 若 f 为非负可测函数，则定义 f 在 Ω 上对 μ 的积分

$$\int_\Omega f\,d\mu := \sup\left\{\int_\Omega h\,d\mu:\ 0\leqslant h\leqslant f,\quad h\text{为}\mathcal{F}-\text{简单函数}\right\}.$$

3) 若 f 为实可测函数，则定义 f 在 Ω 上对 μ 的积分

$$\int_\Omega f\,d\mu := \int_\Omega f^+\,d\mu - \int_\Omega f^-\,d\mu.$$

若 $\int_\Omega f^+\,d\mu$, $\int_\Omega f^-\,d\mu$ 都是 $+\infty$, 则称 f 在 Ω 上对 μ 的 **积分不存在**. 若 $\int_\Omega f^+\,d\mu$, $\int_\Omega f^-\,d\mu$ 有一有限，则称 f 在 Ω 上对 μ 的 **积分存在**.

4) 若 f 为复可测函数，$f=f_1+if_2$, 且 $\int_\Omega f_k\,d\mu$, $k=1,2$, 存在，则定义 f 在 Ω 上对 μ 的积分

$$\int_\Omega f\,d\mu := \int_\Omega f_1\,d\mu + i\int_\Omega f_2\,d\mu$$

若 $\int_\Omega f\,d\mu$ 有限，则称 f 对 μ **可积**.

5) 当 μ 为概率测度时，则称 f 在 Ω 上对 μ 的积分为 f **对** μ **的 (数学) 期望**, 当 f 可积时，则称 **数学期望有限**, 当 $\int_\Omega f\,d\mu$ 存在时，则称 **期望存在**(与本科教科书的规定略有不同).

6) 若 $F\in\mathcal{F}$, 则称 $\int_\Omega fI_F\,d\mu$ 为 f **在** F **上对** μ **的积分**.

f 在 Ω 上对 μ 的积分常简称为 f **对** μ **的积分** (当积分区域 Ω 不致混淆时), 或称 f **的积分** (当 μ 不混淆时), 而 $\int_\Omega f\,d\mu$ 有时写成 $\int_\Omega f(\omega)\,\mu(d\omega)$, 或简记作 $\int_\Omega f$, $\int f\,d\mu$ 或 $\int f$. f 对概率测度 μ 的积分常记作 $\mathbf{E}_\mu f$ 或 $\mathbf{E}f$.

f **在集合** F **上的积分** 是 f 在 F 上对 μ 的积分的简称 (当 μ 不致混淆时). 当 μ 为概率测度时，则 $\int_F f\,d\mu$ 记作 $\mathbf{E}(fI_F)$ 或 $\mathbf{E}[f;F]$.

由上述定义可以看出还须证明 $\mathcal{F}$–简单函数的积分定义的合理性 (由它直接得到 2)-4) 的合理性). 这将由下面的引理完成.

3. 引理 设 f 是 $(\Omega,\mathcal{F},\mu)$ 上的非负 $\mathcal{F}-$ 简单函数，则 $\int f$ 的定义是合理的，即唯一确定，且具有下列性质：

$$f\leqslant g \implies \int f\leqslant\int g; \tag{1}$$

$$c\geqslant 0 \implies \int (cf)=c\int f; \tag{2}$$

$$\varphi(A):=\int_A f \text{ 在}\mathcal{F}\text{上是}\sigma-\text{可加的}. \tag{3}$$

因而当 f 为非负、实、复值可测函数时，$\int f$ 的定义是合理的.

证明 设 $A_i\in\mathcal{F}$, $i=1,2,\cdots,m$, 两两不交且 $\bigcup_{i=1}^{\infty}A_i=\Omega$, $B_j\in\mathcal{F}$, $j=1,2,\cdots,n$, 两两不交且 $\bigcup_{j=1}^{n}B_j=\Omega$,

$$f=\sum_{i=1}^{m}x_iI_{A_i}=\sum_{j=1}^{n}y_jI_{B_j},$$

则当 $A_i\cap B_j\neq\varnothing$ 时，$x_i=y_j$ 当 $A_i\cap B_j=\varnothing$ 时，$\mu(A_i\cap B_j)=0$, 于是

$$\begin{aligned}\sum_{i=1}^{m}x_i\mu(A_i)&=\sum_{i=1}^{m}\sum_{j=1}^{n}x_i\mu(A_i\cap B_j)\\&=\sum_{i=1}^{m}\sum_{j=1}^{n}y_j\mu(A_i\cap B_j)\\&=\sum_{j=1}^{n}y_j\sum_{i=1}^{m}\mu(A_i\cap B_j)\\&=\sum_{j=1}^{n}y_j\mu(B_j).\end{aligned}$$

故定义 2(1) 规定的 $\int f$ 与 f 的表达式无关. 因而定义 2(1) 合理.

其次，为证 (1), 设

$$f=\sum_{i=1}^{m}x_iI_{A_i},\qquad g=\sum_{j=1}^{n}y_jI_{B_j} \tag{4}$$

则和上面证明类似，当 $A_i\cap B_j\neq\varnothing$ 时，$x_i\leqslant y_j$；当 $A_i\cap B_j=\varnothing$ 时，$\mu(A_i\cap B_j)=0$. 于是

$$\int f=\sum_{i=1}^{m}\sum_{j=1}^{n}x_i\mu(A_i\cap B_j)\leqslant\sum_{j=1}^{n}\sum_{i=1}^{m}y_j\mu(A_i\cap B_j)=\int g.$$

(2) 由

$$\int(cf)=\sum_{i=1}^{m}cx_i\mu(A_i)=c\sum_{i=1}^{m}x_i\mu(A_i)=c\int f.$$

即得. 又由 $fI_A=\sum_{i=1}^{m}x_iI_{A\cap A_i}+0I_{A^c}$ 知

$$\varphi(A)=\int_A f\,d\mu=\int fI_A\,d\mu=\sum_{i=1}^{m}x_i\mu(A\cap A_i).$$

因为 μ 为 $\sigma-$ 可加，所以 $\mu(\cdot\cap A_i)$ 也 $\sigma-$ 可加，从而 $\varphi(A)$ $\sigma-$ 可加，故 (3) 获证. □

4. 引理 设 f,g 为非负可测函数，且 $g\leqslant f$, 则 $\int g\leqslant\int f$.

证明 对任何满足 $0\leqslant h\leqslant g$ 的 $\mathscr{F}-$ 简单函数 h 来说，必有 $0\leqslant h\leqslant f$. 因而

$$\sup\left\{\int h:\ 0\leqslant h\leqslant g,\ h\text{为}\mathscr{F}-\text{简单函数}\right\}$$
$$\leqslant\sup\left\{\int h:\ 0\leqslant h\leqslant f,\ h\text{为}\mathscr{F}-\text{简单函数}\right\}$$

这就是所要证的. □

5. 定理 (单调收敛定理) 若 $f_n, n\in\mathbb{N}$, 是非负可测函数列，且 $f_n\uparrow f$, 则

$$\lim_{n\to+\infty}\int f_n=\int f.$$

证明 由引理 4 知 $\int f_n$ 为非降数列，因而 $\lim_{n\to\infty}\int f_n$ 存在 (可能是 $+\infty$). 记作 a. 又由 $f_n \leqslant f$, 对一切 $n \in \mathbb{N}$ 成立，所以有 $a \leqslant \int f$. 往证：

$$(1) \qquad a \geqslant \int f.$$

对任意满足 $0 \leqslant h \leqslant f$ 的 $\mathcal{F}-$ 简单函数 h, 设

$$h = \sum_{j=1}^{m} x_j I_{A_j} + (+\infty) I_{A_{m+1}}, \qquad j = 1, 2, \cdots, m.$$

任取 $c \in (0,1)$, $k \in \mathbb{N}$, 令

$$h_{c,k} := \sum_{j=1}^{m} (cx_j) I_{A_j} + k I_{A_{m+1}},$$
$$\Omega_n := \{\omega \in \Omega: \ h_{c,k}(\omega) \leqslant f_n(\omega)\}.$$

当然 Ω_n 与 c, k 有关，当 c, k 给定， $n \to \infty$ 时， $\Omega_n \uparrow \Omega$. 由

$$f_n \geqslant f_n I_{\Omega_n} \geqslant h_{c,k} I_{\Omega_n}$$

及引理 4 知

$$\int f_n \geqslant \int h_{c,k} I_{\Omega_n} = \int_{\Omega_n} h_{c,k}.$$

由引理 3(3) 知 $\varphi(A) = \int_A h_{c,k}$ 非负且 $\sigma-$ 可加 (即为测度), 因而 φ 下连续，从而令 $n \to \infty$ 得到

$$a \geqslant \lim_{n\to\infty} \int_{\Omega_n} h_{c,k} = \int_{\Omega} h_{c,k} = \sum_{j=1}^{m} (cx_j)\mu(A_j) + k\mu(A_{m+1}).$$

再令 $c \to 1$, $k \to \infty$ 即得

$$a \geqslant \sum_{j=1}^{m} x_j \mu(A_j) + \infty I_{A_{m+1}} = \int h.$$

由定义 2.2) 知 (1) 成立. 故定理获证. □

6. 推论 对任何以非负 $\mathcal{F}-$ 可测函数 f 为极限的非负不降 $\mathcal{F}-$ 简单函数列 f_n, $n\in\mathbb{N}$, 来说都有

$$\lim_n \int f_n = \int f.$$

证明 由于 $0\leqslant f_n\uparrow f$. 所以由定理 5 上式成立.

注 推论 6 实际上给出了非负可测函数的积分的另一定义. 而且它在讨论非负可测函数积分的性质时很有用.

7. 定义 给定 $(\Omega,\mathcal{F},\mu)$. 设有一与 $\omega\in\Omega$ 有关的性质 $\mathcal{P}=\{\mathcal{P}(\omega):\ \omega\in\Omega\}$. 若

$$\{\omega\in\Omega:\ \mathcal{P}(\omega)\text{不成立}\}$$

是 $\mu-$ 零集, 则称 $\mathcal{P}$ **对几乎一切** $\omega\in\Omega$ **成立**, 简记作 $\mathcal{P}$ 成立, μ-a.e.. 例如, 设 f,g 都是 Ω 到 $\overline{\mathbb{R}}$ 的函数, 若有 $\{\omega\in\Omega:\ f(\omega)\neq g(\omega)\}$ 是一 $\mu-$ 零集, 则称 f **与** g **几乎处处相等**, 简记作 $f=g$, $\mu-$a.e.; 若 $\{\omega\in\Omega:\ f(\omega)\leqslant g(\omega)\}$ 是一 $\mu-$ 零集, 则称 f **几乎处处大于** g, 简记作 $f>g$, $\mu-$a.e.. 又如, 若 $f:\ \Delta\mapsto\overline{\mathbb{R}}$, 且 Δ^c 为一 $\mu-$ 零集, 则称 f **为几乎处处有定义**, 简记作 f $\mu-$a.e. 有定义.

当 μ 不致混淆时, "$\mu-$a.e." 简记作 "a.e.".

当 μ 为概率测度时, 常称 "几乎处处" 为 "几乎必然", 简记作 "a.s.".

8. 引理 给定 $(\Omega,\mathcal{F},\mu)$, 若 f,g 都是 $\mathcal{F}-$ 可测函数, 且 $f=g$, $\mu-$a.e., 则 $\int f=\int g$.

证明 因为由定义 7, $N:=\{\omega:f(\omega)\neq g(\omega)\}$ 是一 $\mu-$ 零集, 且 $N\in\mathcal{F}$, 所以 $\mu(N)=0$, 因而 $fI_{N^c}=gI_{N^c}$, 若 $f=\sum_{k=1}^m x_kI_{A_k}$, $g=\sum_{\ell=1}^n y_\ell I_{B_\ell}$, 为非负 $\mathcal{F}-$ 简单函数, 则在 $\mu(A_k\cap B_\ell)\neq 0$ 时,

$x_k = y_\ell$, 从而

$$\int f\,d\mu = \sum_{\substack{k=1 \\ \mu(A_k\cap B_\ell)\neq 0}}^{m}\sum_{\ell=1}^{n} x_k\mu(A_k\cap B_\ell)$$
$$= \sum_{\substack{k=1 \\ \mu(A_k\cap B_\ell)\neq 0}}^{m}\sum_{\ell=1}^{n} y_\ell\mu(A_k\cap B_\ell) = \int g\,d\mu.$$

若 f,g 为非负 $\mathcal{F}-$ 可测函数, 则存在 $\mathcal{F}-$ 简单函数列 $\{f_n\}$, $\{g_n\}$, 使 $0\leqslant f_n\uparrow f$, $0\leqslant g_n\uparrow g$. 于是

$$0\leqslant g_n' := f_nI_{N^c} + g_nI_N \uparrow fI_{N^c} + gI_N = g$$

且 g_n' 为 $\mathcal{F}-$ 简单函数, $f_n = g_n'$ a.e.. 故

$$\int f = \lim_{n\to\infty}\int f_n = \lim_{n\to\infty}\int g_n' = \int g.$$

若 f,g 为实可测函数, $f=g$ a.e., 则必有 $f^{\pm}=g^{\pm}$ a.e., 因而

$$\int f^{\pm} = \int g^{\pm}, \qquad \int f = \int g.$$

若 f,g 为复可测函数, $f=g$ a.e.. 设 $f=f_1+if_2$, $g=g_1+ig_2$, 则必有 $f_i=g_i$ a.e., $i=1,2$, 于是

$$\int f_i = \int g_i, \quad i=1,2, \quad 因而\int f = \int g.$$

9. 定义 若 f a.e. 可测 (即存在可测函数 f' 使得 $f=f'$ a.e.)(或 a.e. 有定义) $f=g$ a.e., g 可测, $\int g$ 存在, 则定义

$$\int_\Omega f\,d\mu := \int_\Omega g\,d\mu.$$

10. 注 若将本节中的所有可测函数改为 a.e. 可测, 则本节中所有关于积分的结论正确 (为什么?)

以后讨论积分时, 当出现的函数是 a.e. 可测时, 所有结论正确.

习题

1. 给出非负可测函数积分的另一种定义：

1) 按照定义 2.1 中定义非负简单函数的积分，证明定义的合理性；

2) 证明非负简单函数的积分具有性质：若 $f \leqslant g$，则 $\int f \leqslant \int g$；

3) 如下定义非负可测函数的积分：若 f 非负可测，$\{f_n\}$ 为简单函数列，满足 $0 \leqslant f_n \uparrow f$，则令 $\int f = \lim\limits_{n\to\infty} \int f_n$；

4) 证明定义 3) 的合理性；

5) 证明单调收敛定理.

2. 证明注 10.

§2. 积分的性质

1. 引理 给定 $(\Omega, \mathcal{F}, \mu)$，则 $\mathcal{F}$- 可测函数对 μ 的积分具有下列的线性性质：

1) 若 $\int f$, $\int g$ 存在且 $\int f + \int g$ 有意义，则 $f+g$ a.e. 有定义，a.e. 可测，$\int (f+g)$ 存在且

$$\int (f+g) = \int f + \int g;$$

2) 若 $\int f$ 存在，则对一切 $A \in \mathcal{F}$，$\int_A f$ 存在，且当 $A \cap B = \varnothing$ 时，$\int_{A\cup B} f = \int_A f + \int_B f$；

3) 若 c 为有限复数，f 为实函数且 $\int f$ 存在或为可积复函数，则

$$\int (cf) = c \int f;$$

4) $\int \bar{f} = \overline{\int f}$，其中 $\bar{z}$ 表示复数 z 的共轭.

证明 1) 的证明 按照积分的定义, 依次由简单情形到一般情形验证 1).

设 f, g 为非负 $\mathscr{F}$– 简单函数

$$f = \sum_{k=1}^{m} x_k I_{A_k}, \qquad g = \sum_{k=1}^{n} y_k I_{B_\ell}.$$

则可写出

$$f + g = \sum_{k=1}^{m} \sum_{\ell=1}^{n} (x_k + y_\ell) I_{A_k \cap B_\ell}.$$

于是,

$$\begin{aligned} \int f + g &= \sum_{k=1}^{m} \sum_{\ell=1}^{n} (x_k + y_\ell) \mu(A_k \cap B_\ell) \\ &= \sum_{k=1}^{m} x_k \sum_{\ell=1}^{n} \mu(A_k \cap B_\ell) + \sum_{\ell=1}^{n} y_\ell \sum_{k=1}^{m} \mu(A_k \cap B_\ell) \qquad (1) \\ &= \int f + \int g. \end{aligned}$$

其次, 设 f, g 为非负 $\mathscr{F}$– 可测函数, 则由定理 4.2.4 知存在 $f_n, g_n, n \in \mathbb{N}$, 都是 $\mathscr{F}$–简单函数, $0 \leqslant f_n \uparrow f$, $0 \leqslant g_n \uparrow g$, 因而 $0 \leqslant f_n + g_n \uparrow f + g$. 由推论 1.6 及 (1) 知

$$\begin{aligned} \int f + \int g &= \lim_{n\to\infty} \int f_n + \lim_{n\to\infty} \int g_n \\ &= \lim_{n\to\infty} \left(\int f_n + \int g_n \right) \qquad (2) \\ &= \lim_{n\to\infty} \int (f_n + g_n) = \int (f + g). \end{aligned}$$

第三, 设 f, g 为实可测函数, 由假设知

$$\int f + \int g = \left(\int f^+ - \int f^- \right) + \left(\int g^+ - \int g^- \right)$$

有意义. 因而 $\int f^+ + \int g^+$ 与 $\int f^- + \int g^-$ 至少有一个有限, 设

(3) $$\int f^- + \int g^- < \infty.$$

由于非负可测函数 $\tilde{f}$ 的积分若有限, 则必有 $\mu(\{\tilde{f} = \infty\}) = 0$ (即 $\tilde{f}$ 几乎处处有限, 记作 $\tilde{f}$ a.e. 有限). 否则令 $h = I_{\{\tilde{f}=\infty\}}$, 则由定义 1.2.2) 知 $\int \tilde{f} \geqslant \int h = \infty \cdot \mu(\{\tilde{f} = \infty\}) = \infty$, 矛盾. 所以 f^-, g^- a.e. 有限. 因而 $f + g$ a.e. 有定义. 又有

(4) $(f+g)^- = [-(f+g)]\vee 0 \leqslant [(-f)\vee 0]+[(-g)\vee 0] = f^- + g^-$, a.e.

于是 $(f + g)^-$ a.e. 有限, a.e. 可测. 因此, 由

$$(f + g)^+ - (f + g)^- = f + g = (f^+ - f^-) + (g^+ - g^-), \qquad \text{a.e.},$$

即得

(5) $$(f + g)^+ + f^- + g^- = (f + g)^- + f^+ + g^+ \qquad \text{a.e.},$$

再由 (2) 得

(6) $$\int (f + g)^+ + \int f^- + \int g^- = \int (f + g)^- + \int f^+ + \int g^+.$$

而由 (3),(4) 知 $\int (f + g)^- < \infty$, 故得

(7) $$\int f + \int g = \int (f + g).$$

同法可证: 当 $\int f^+ + \int g^+$ 有限时, (7) 仍然成立.

对于 f, g 为复函数的情形, 应用实函数的结论分实部, 虚部讨论即可得出结论.

2) 的证明 $\int f$ 存在, 就有 $\int f^+$, $\int f^-$ 有一有限, 设 $\int f^- < \infty$, 则 $\int (fI_A)^-$, $\int (fI_B)^-$ 都有限, 因而, $\int (fI_A)$, $\int (fI_B)$, $\int (fI_A) + \int (fI_B)$ 存在. 于是由 1) 及 $A \cap B = \varnothing$ 知

$$\begin{aligned}\int_A f + \int_B f &= \int fI_A + \int fI_B = \int (fI_A + fI_B) \\ &= \int fI_{A\cup B} = \int_{A\cup B} f.\end{aligned}$$

3) 的证明：先考虑 $c \geqslant 0$ 情形. 当 f 为非负 $\mathcal{F}$– 简单函数时

$$\int (cf) = c\int f, \tag{8}$$

即引理 1.3 的 (2) 式. 当 f 非负 $\mathcal{F}$– 可测时，cf 亦然. 取 $0 \leqslant f_n \uparrow f$. f_n 为 $\mathcal{F}$– 简单函数. 则 $0 \leqslant cf_n \uparrow cf$. 于是由推论 1.6 知

$$\int (cf) = \lim_{n\to\infty} \int (cf_n) = \lim_{n\to\infty} c\int f_n = c\int f.$$

当 f 为实可测函数且 $\int f$ 存在时，则有

$$\begin{aligned} c\int f &= c\int f^+ - c\int f^- = \int cf^+ - \int cf^- \\ &= \int (cf)^+ - \int (cf)^- = \int cf. \end{aligned}$$

故当 $c \geqslant 0$, f 为 $\int f$ 存在的实函数时 $\int cf$ 存在，(8) 成立.

当 $c < 0$ 时，f 为实函数且 $\int f$ 存在时，

$$\begin{aligned} c\int f &= -(-c)\int f = (-c)\left[\int f^- - \int f^-\right] \\ &= \int (-c)f^- - \int (-c)f^+ = \int (cf)^+ - \int (cf)^- = \int (cf), \end{aligned}$$

即 (8) 成立.

当 $c = c_1 + ic_2$, f 为积分存在的实函数时，

$$\begin{aligned} c\int f &= c_1\int f + ic_2\int f \\ &= \int (c_1 f) + i\int (c_2 f) \\ &= \int (c_1 + ic_2)f = \int (cf). \end{aligned}$$

设 f 为复函数且可积，c 为复数，记

$$f := f_1 + if_2, \qquad c = c_1 + ic_2,$$

则 $\int f_1$、$\int f_2$ 都有限，因此由 (8)、(7) 立得：

$$\begin{aligned}
c\int f &= (c_1 + ic_2)\Big(\int f_1 + i\int f_2\Big)\\
&= \Big(c_1\int f_1 - c_2\int f_2\Big) + i\Big(c_1\int f_2 + c_2\int f_1\Big)\\
&= \int (c_1 f_1 - c_2 f_2) + i\int (c_1 f_2 + c_2 f_1)\\
&= \int [(c_1 f_1 - c_2 f_2) + i(c_1 f_2 + c_2 f_1)] = \int (cf).
\end{aligned}$$

4) 的证明 设 $f = f_1 + if_2$，则由复函数积分的定义知

$$\begin{aligned}
\int \bar{f} &= \int (f_1 - if_2) = \int f_1 + i\int (-f_2)\\
&= \int f_1 - i\int f_2 = \overline{\int f_1 + i\int f_2} = \overline{\int f}. \quad \square
\end{aligned}$$

2. 推论 若 $\int f$, $\int g$ 存在，a、b 为有限复数，且 $a\int f + b\int g$ 有意义，则 $\int (af + bg)$ 存在，且

$$\int (af + bg) = a\int f + b\int g.$$

证明 由 $a\int f + b\int g$ 有意义知 $a\int f$、$b\int g$ 有意义，因而，由引理 1 的 1) 及 3) 即得结论． $\square$

3. 引理 给定 $(\Omega, \mathcal{F}, \mu)$，则 $\mathcal{F}$– 可测函数的积分有下列序性质：

1) 若 f, g 为实函数，$\int f$、$\int g$ 存在，若 $f \geqslant g$ a.e.，则

$$\int_A f \geqslant \int_A g, \qquad \forall A \in \mathcal{F}; \tag{1}$$

反之，若 μ 为 σ– 有限测度，则逆命题成立；

2) 若 $\int f$ 存在，则 $|\int f| \leqslant \int |f|$；

3) $f \geqslant 0$，则 $\int f = 0$ 的充要条件是 $f = 0$ a.e..

证明 1) 由引理 1.8 知，只须证 $f \geqslant g$ 的情形.

当 f, g 非负时，由引理 1.4 及 $fI_A \geqslant gI_A$ 知 (1) 成立. 当 f、g 为实可测函数时，由 $f \geqslant g$ 知 $f^+ \geqslant g^+$, $f^- \leqslant g^-$, 因而

$$\int_A f^+ \geqslant \int_A g^+, \quad \int_A f^- \leqslant \int_A g^-, \qquad \forall A \in \mathcal{F}.$$

故

$$\int_A f = \int_A f^+ - \int_A f^- \geqslant \int_A g^+ - \int_A g^+ = \int_A g, \quad \forall A \in \mathcal{F}.$$

反之，若 μ 为 $\sigma-$ 有限，且对一切 $A \in \mathcal{F}$、$\int_A f\,d\mu \geqslant \int_A g\,d\mu$，往证 $f \geqslant g$ a.e.，即 $\mu(\{f < g\}) = 0$. 如若不然，令

$$A_{ab} := \{f \leqslant a < b \leqslant g\}, \qquad a, b\text{为有理数，且}a < b.$$

则

$$\{f < g\} = \bigcup_{\substack{a<b \\ a,b\text{ 为有理数}}} A_{ab}.$$

因而存在 a, b 为有理数，$a < b$, 使 $\mu(A_{ab}) > 0$. 若 $\mu(A_{ab}) < \infty$, 则

$$\int_{A_{ab}} f \leqslant \int fI_{A_{ab}} \leqslant a\mu(A_{ab}) < b\mu(A_{ab}) \leqslant \int_{A_{ab}} g.$$

这与前提矛盾. 若 $\mu(A_{ab}) = \infty$、则由 μ 为 $\sigma-$ 有限知存在 $\Omega_n \in \mathcal{F}$, $n \in \mathbb{N}$, 两两不交，$\bigcup_{n=1}^{\infty} \Omega_n = \Omega$、$\mu(\Omega_n) < \infty$. 因而存在 m, 使 $0 < \mu(A_{ab} \cap \Omega_m) < \infty$, 于是

$$\int_{A_{ab}\cap\Omega_m} f \leqslant a\mu(A_{ab} \cap \Omega_m) < b\mu(A_{ab} \cap \Omega_m) \leqslant \int_{A_{ab}\cap\Omega_m} g,$$

同样与前提矛盾. 故 1) 获证.

2) 当 f 为实函数时，

$$\int |f| = \int f^+ + \int f^- \geqslant \left|\int f^+ - \int f^-\right| = \left|\int f\right|.$$

当 f 为复函数时，若 $\int|f|=\infty$. 不等式自然成立，若 $\int|f|<\infty$, 设 $f=f_1+if_2$. 则由上面已经证明的事实知 $\int f_i, i=1,2$ 有限，因而 $\int f$ 有限．于是可设

$$\int f = re^{it},\quad r,t\in\mathbb{R}.\qquad 因而\, r=\left|\int f\right|.$$

再设 $f(\omega)=\rho(\omega)e^{i\alpha(\omega)}$, 其中 $\rho(\omega)=|f(\omega)|$, $\alpha(\omega)$ 为 $f(\omega)$ 的幅角 (当 $\rho(\omega)=0$ 时，取 $\alpha(\omega)=0$). 则易证 $\rho(\cdot),\alpha(\cdot)$ 均为 $\mathcal{F}$ 可测函数，于是

$$\begin{aligned}\left|\int f\right| &= r = e^{-it}\int f = e^{-it}\int \rho e^{i\alpha} = \int \rho e^{i(\alpha-t)}\\ &= \int \rho\cos(\alpha-t)\qquad \left(因为\int \rho e^{i(\alpha-t)} = r\in\mathbb{R}\right)\\ &\leqslant \int \rho = \int |f|.\end{aligned}$$

故 2) 获证．

3) 由引理 1.8, 知 $f=0$ a.e. $\Longrightarrow \int f\,d\mu=0$. 反之，若 $f\geqslant 0$ a.e. 且 $\int f=0$, 则必有 $f=0$ a.e.. 否则 $\mu(\{f>0\})>0$. 而由 $\{f>0\}=\bigcup_{n=1}^{\infty}\{f>\frac{1}{n}\}$ 及 μ 的下连续性

$$\mu(\{f>0\})=\lim_{n\to\infty}\mu\left(\left\{f>\frac{1}{n}\right\}\right),$$

因而存在 $n\in\mathbb{N}$, 使 $\mu(\{f>\frac{1}{n}\})>0$. 于是

$$\int f\geqslant\int \frac{1}{n}I_{\{f>\frac{1}{n}\}}=\frac{1}{n}\mu\left(\left\{f>\frac{1}{n}\right\}\right)>0,$$

与假设矛盾．故 3) 获证．因而引理证毕． □

现在来讨论积分的可积性质．

4. 引理 给定 $(\Omega,\mathcal{F},\mu)$, 则 $\mathcal{F}$ 可测函数具有下列可积性质：

1) f 可积的充要条件是 f a.e. 可测且 $\int|f|<\infty$, 当 f 可积时，f a.e. 有限；

2) 若 $|f| \leqslant g$, g 可积, f 可测, 则 f 可积;

3) 若 f, g 可积, 则 $f+g$ 可积.

证明 1) 设 $f = f_1 + if_2 = (f_1^+ - f_1^-) + i(f_2^+ - f_2^-)$. 若 f 可测, $\int |f| < \infty$. 则 $f_i^{\pm}$, $i=1,2$, 都可测, 且 $f_i^{\pm} \leqslant |f|$, 因而由引理 3.1) 知 $\int f_i^{\pm} < \infty$, $i=1,2$. 故 $\int f$ 有限, 即 f 可积.

若 f 可积, 则 f a.e. 可测, $\int f$ 有限, 因而 $\int f_i^{\pm}$, $i=1,2$, 都有限, 而

$$|f| \leqslant |f_1| + |f_2| = (f_1^+ + f_1^-) + (f_2^+ + f_2^-),$$

故由引理 3.1) 及引理 1.1) 知

$$\int |f| \leqslant \int f_1^+ + \int f_1^- + \int f_2^+ + \int f_2^- < \infty.$$

此外, 当 f 可积时, $\int |f| < \infty$, 于是

$$\infty \cdot \mu(\{|f| = \infty\}) = \int \infty \cdot I_{\{|f| = \infty\}} \leqslant \int |f| < \infty,$$

故 $\mu(\{|f| = \infty\}) = 0$, 即 f a.e. 有限.

2) 由 f 可测知 $|f|$ 可测, 于是由引理 3.1) 知 $\int |f| \leqslant \int g < \infty$. 故由 1) 知 f 可积.

3) 当 f, g 可积时, 由 1) 知 f, g 都可测且 a.e. 有限, 因而 $f+g$ a.e. 可测, 且 $|f+g| \leqslant |f| + |g|$. 而由 1) 知:

$$\int (|f| + |g|) = \int |f| + \int |g| < \infty,$$

故 $|f| + |g|$ 可积, 因而由 2) $f+g$ 可积. □

5. 引理 **(Schwarz 不等式)**. 给定 $(\Omega, \mathcal{F}, \mu)$, f, g $\mathcal{F}$−a.e. 可测, 则 $(\int |fg|)^2 \leqslant \int |f|^2 \cdot \int |g|^2$.

证明 若 $f=0$ a.e. 或 $g=0$ a.e., 则命题显然成立. 于是不妨设 $\int|f|^2>0$ $\int|g|^2>0$. 此时若 $\int|f|^2=\infty$. 或 $\int|g|^2=\infty$, 则命题也显然成立. 故只需考察 $0<\int|f|^2, \int|g|^2<\infty$ 的情形. 由

$$0\leqslant\int(x|f|+y|g|)^2=\left(\int|f|^2\right)x^2+2\left(\int|fg|\right)xy+\left(\int|g|^2\right)y^2$$

对所有 $x,y\in\mathbb{R}$ 成立, 即得

$$\left(\int|f||g|\right)^2\leqslant\int|f|^2\cdot\int|g|^2\quad\Box$$

习题

1. 设 μ_1、μ_2 是可测空间 $(\Omega,\mathcal{F})$ 上的测度, a_1、a_2 是两个非负有限数, $\mu=a_1\mu_1+a_2\mu_2$. 试证: 若 f 对 μ_1,μ_2 的积分存在且 $a_1\int fd\mu_1+a_2\int fd\mu_2$ 有意义, 则 f 对 μ 的积分也存在, 且

$$\int fd\mu=a_1\int fd\mu_1+a_2\int fd\mu_2.$$

2. (积分中值定理) 设 f,g 是测度空间 $(\Omega,\mathcal{F},\mu)$ 上的可测函数, g 对 μ 可积, $-\infty<a\leqslant f\leqslant b<\infty$, a.e., 则存在一个常数 $c\in[a,b]$, 使

$$\int_\Omega f|g|d\mu=c\int_\Omega|g|d\mu.$$

特别, 若 μ 有限, 则 $\int fd\mu=c\mu(\Omega)$.

3. 设 f,g 是 $(\Omega,\mathcal{F},\mu)$ 上取有限值的 $\mathcal{F}$ 简单函数, 若 f,g 之一对 μ 可积, 则 fg 也可积.

4. 设 f 是 $(\Omega,\mathcal{F},\mu)$ 上的可积函数, 则对每一正数 ϵ, $\mu(\{\omega:|f(\omega)|\geqslant\epsilon\})<\infty$.

5. 设 Ω 是全体正整数组成的空间, $\mathcal{F}$ 是 Ω 的一切子集作成的 σ– 代数, 对于 $A\in\mathcal{F}$, $\mu(A)=A$ 中点的个数, 则 $(\Omega,\mathcal{F},\mu)$ 是一个测度空间. 称此 μ 为计数测度. 讨论此空间上的函数的积分存在的充分与必要条件.

6. 设 f, g 是 $(\Omega, \mathcal{F})$ 上的可测函数，$\mathcal{P}(\mathcal{F})$ 表示 $\mathcal{F}$ 上的一切概率测度的集. 若对一切 $\nu \in \mathcal{P}(\mathcal{F})$, $\int f d\nu = \int g d\nu$, 则 $f(\omega) = g(\omega)$, 对一切 $\omega \in \Omega$ 成立.

§3. 期望的性质及 L-S 积分表示

在前面已经定义概率空间 $(\Omega, \mathcal{F}, \mathbf{P})$ 上的实值 (或复值) 可测函数为随机变量 (r.v.). 而随机变量对 $\mathbf{P}$ 的积分就是它的数学期望. 这样就可以将概率论严格地建立在测度论的基础上. 这一节将给出随机变量的期望, 方差的初等性质的证明, 给出随机变量的函数的概率分布及其期望的 L-S 积分表示，作为数学工具. 将介绍独立事件类扩张定理及积分变换定理.

1. 引理 给定概率空间 $(\Omega, \mathcal{F}, \mathbf{P})$, 设 $X_1, \cdots, X_n$ 为其上的 r.v. , $c_k, k = 1, 2, \cdots, n$, 为有限常数 (实或复数). 若 $\sum_{k=1}^n c_k \mathbf{E} X_k$ 有意义，则 $\mathbf{E}[\sum_{k=1}^n c_k X_k]$ 存在，且

$$\mathbf{E}\left[\sum_{k=1}^n c_k X_k\right] = \sum_{k=1}^n c_k \mathbf{E} X_k.$$

此定理由 $\mathbf{P}$ 的积分的线性性质立即得到.

为了讨论期望的乘法性质，我们引进概率论中的一个基本概念—随机变量独立的概念.

2. 定义 设 X_k 是 $(\Omega, \mathcal{F}, \mathbf{P})$ 上的 $m_k(\in \mathbb{N})$ 维实 r.v. , $k = 1, \cdots, n$, 若对一切 $x_k \in \overline{\mathbb{R}}^{m_k}$, 有

$$\mathbf{P}\Big(\bigcap_{k=1}^m \{X_k \leqslant x_k\}\Big) = \prod_{k=1}^n \mathbf{P}(X_k \leqslant x_k). \tag{1}$$

则称 $X_1, \cdots, X_n$ **独立**.

设 $X_k = X_k' + iX_k''$ 是 m_k 维复 r.v. , $k = 1, 2, \cdots, n$, 如果 $(X_1', X_1''), \cdots, (X_n', X_n'')$ 独立，则称 $X_1, \cdots, X_n$ **独立**.

3. 定理 (独立事件类的扩张定理) $(\Omega, \mathcal{F}, \mathbf{P})$ 为概率空间, $\mathcal{C}_k \subset \mathcal{F}$. 为 $\pi-$ 系, $\Omega \in \mathcal{C}_k$, $k = 1, 2, \cdots, n$, 且对一切 $A_k \in \mathcal{C}_k$, $k = 1, 2, \cdots, n$, 有

$$\mathbf{P}\Big(\bigcap_{k=1}^{n} A_k\Big) = \prod_{k=1}^{n} \mathbf{P}(A_k), \tag{1}$$

则 (1) 对一切 $A_k \in \sigma(\mathcal{C}_k)$, $k = 1, 2, \cdots, n$, 成立.

证明 1) 对一给定的 $\ell \in \mathbb{N} \cap [1, n]$, 令

$$\Lambda_\ell := \{A_\ell \in \sigma(\mathcal{C}_\ell) : \ \forall\ A_k \in \mathcal{C}_k, \quad k \neq \ell, \quad (1)\text{成立}\}$$

显然有 $\mathcal{C}_\ell \subset \Lambda_\ell \subset \sigma(\mathcal{C}_\ell)$. 今往证 Λ_ℓ 为 $\mathcal{F}$ 中的 $\lambda-$ 系, 由假设知 $\Omega \in \mathcal{C}_\ell$. 若 $A, B \in \Lambda_\ell$, $A \supset B$, 则对一切 $A_k \in \mathcal{C}_k$, $k \neq \ell$, 由 Λ_ℓ 的定义知

$$\begin{aligned}
&\mathbf{P}\Big(\Big(\bigcap_{\substack{k=1\\k\neq\ell}}^{n} A_k\Big)\bigcap(A \setminus B)\Big)\\
&\qquad= \mathbf{P}\Big(\Big(\bigcap_{\substack{k=1\\k\neq\ell}}^{n} A_k\Big)\bigcap A\Big) - \mathbf{P}\Big(\Big(\bigcap_{\substack{k=1\\k\neq\ell}}^{n} A_k\Big)\bigcap B\Big)\\
&\qquad= \prod_{\substack{k=1\\k\neq\ell}}^{n} \mathbf{P}(A_k)\mathbf{P}(A) - \prod_{\substack{k=1\\k\neq\ell}}^{n} \mathbf{P}(A_k)\mathbf{P}(B)\\
&\qquad= \prod_{\substack{k=1\\k\neq\ell}}^{n} \mathbf{P}(A_k)\mathbf{P}(A \setminus B).
\end{aligned}$$

故 Λ_ℓ 对真差封闭, 其次设 $A_{\ell_m} \in \Lambda_\ell$, $m \in \mathbb{N}$, $A_{\ell_m} \uparrow A_\ell (m \to \infty)$. 则由 $\mathbf{P}$ 的下连续性及 Λ_ℓ 的定义,

$$\begin{aligned}
\mathbf{P}\Big(\bigcap_{k=1}^{n} A_k\Big) &= \lim_{m\to\infty} \mathbf{P}\Big(\Big(\bigcap_{\substack{k=1\\k\neq\ell}}^{n} A_k\Big)\bigcap A_{\ell_m}\Big)\\
&= \lim_{m\to\infty} \prod_{\substack{k=1\\k\neq\ell}}^{n} \mathbf{P}(A_k)\mathbf{P}(A_{\ell_m}) = \prod_{k=1}^{n} \mathbf{P}(A_k).
\end{aligned}$$

故 Λ_ℓ 对不降集列的并封闭, 因而 Λ_ℓ 是一 $\lambda-$ 系. 故由定理 3.1.15 知 $\Lambda_\ell \supset \sigma(\mathcal{C}_\ell)$, 因而 $\Lambda_\ell = \sigma(\mathcal{C}_\ell)$. 即将假设中的任一 $\mathcal{C}_\ell$ 可以换成 $\sigma(\mathcal{C}_\ell)$, (1) 式仍然成立.

2) 由 1) 知先将假设中的 $\mathcal{C}_1$ 换成 $\sigma(\mathcal{C}_1)$, (1) 成立, 由于 $\Omega \in \sigma(\mathcal{C}_1)$ 且仍为 $\pi-$ 系, 所以 $\sigma(\mathcal{C}_1), \mathcal{C}_2, \cdots, \mathcal{C}_n$ 仍然满足定理的假设, 于是再应用 1) 将 $\mathcal{C}_2$ 换成 $\sigma(\mathcal{C}_2)$ 后, (1) 式仍然成立. 依次多次应用 (1), 经过 n 次即得证: (1) 式对一切 $A_k \in \sigma(\mathcal{C}_k)$, $k = 1, 2, \cdots, n$, 成立. □

4. 定理 1) 若 X_k 是 $(\Omega, \mathcal{F}, \mathbf{P})$ 上的 $m_k(\in \mathbb{N})$ 维实 r.v., $k = 1, 2, \cdots, n$, $X_1, \cdots, X_n$ 独立, 则对一切 $B_k \in \mathcal{B}^{m_k}$, $k = 1, 2, \cdots, n$, 有

$$\mathbf{P}\Big(\bigcap_{k=1}^{n}\{X_k \in B_k\}\Big) = \prod_{k=1}^{n} \mathbf{P}(\{X_k \in B_k\}).$$

因而若 $f_k: \mathbb{R}^{m_k} \mapsto \mathbb{R}^{n_k}$ (或 $\mathbb{C}^{n_k}$) 是 Borel 可测函数, 则 $f_k(X_k)$, $k = 1, 2, \cdots, n$, 独立.

2) 若 $X_k = X_k' + iX_k''$ 是 m_k 维复 r.v., $k = 1, 2, \cdots, n$, $X_1, \cdots, X_n$ 独立, $f_k: \mathbb{C}^{m_k} \mapsto \mathbb{C}^{n_k}$ (或 $\mathbb{R}^{n_k}$) Borel 可测函数, 则 $f_k(X_k), k = 1, 2, \cdots, n$, 独立.

证明 令

$$\mathcal{C}_k := \Big\{\{X_k \in (-\infty, x_k]\}: \ x_k \in \mathbb{R}^{m_k}\Big\} \bigcup \{\Omega\}, \qquad k = 1, 2, \cdots, n,$$

则由 $X_1, \cdots, X_k$ 独立的定义知 $\mathcal{C}_k$, $k = 1, 2, \cdots, n$, 满足定理 3 的条件, 而 $\sigma(\mathcal{C}_k) = X^{-1}(\mathcal{B}^{m_k}) = \{\{X_k \in B_k\}: \ B_k \in \mathcal{B}^{m_k}\}$, 故 1) 获证.

对一切 $B_k \in \mathcal{B}^{2m_k}$, $k = 1, 2, \cdots, n$, 由假设有

$$\mathbf{P}\Big(\bigcap_{k=1}^{n}\{(X_k', X_k'') \in B_k\}\Big) = \prod_{k=1}^{n} \mathbf{P}\big(\{(X_k', X_k'') \in B_k\}\big).$$

因而，若令

$$g_k(x_k', x_k'') := f_k(x_k' + \mathrm{i}x_k''), \qquad x_k', x_k'' \in \mathbb{R}^{m_k}, \quad k = 1, 2, \cdots, n,$$

则 g_k 为 $2m_k$ 个实变量的 $\mathscr{B}^{2m_k}$ 可测函数 (实或复值)，且容易看出 $g_k(X_k', X_k'') = f_k(X_k)$，$k = 1, 2, \cdots, n$，独立. □

5. 定理 设 $X_1, \cdots, X_n$ 为独立 (实或复) r.v.，且 $\mathbf{E}X_k$ 有限，$k = 1, 2, \cdots, n$，则

$$\mathbf{E}(X_1 \cdots X_n) = \prod_{k=1}^{n} \mathbf{E}X_k. \tag{1}$$

证明 只需证 $n = 2$ 的情形，对一般情形不难用数学归纳法推出.

1) 设 ξ, η 是独立的非负可测简单函数 (即非负的只取有限个值的随机变量)：

$$\xi = \sum_{j=1}^{m} a_j I_{A_j}, \qquad \eta = \sum_{k=1}^{n} b_k I_{B_k}.$$

这时，

$$\xi\eta = \sum_{j=1}^{m} \sum_{k=1}^{n} a_j b_k I_{A_j} I_{B_k}.$$

由于 ξ, η 独立，$\mathbf{P}(A_j \cap B_k) = \mathbf{P}(A_j)\mathbf{P}(B_k)$，及 ξ, η 非负可测，因而数学期望 $\mathbf{E}\xi\eta$, $\mathbf{E}\xi$, $\mathbf{E}\eta$ 存在，于是利用非负可测简单函数积分的定义

$$\begin{aligned}
\mathbf{E}\xi\eta &= \sum_{j=1}^{m} \sum_{k=1}^{n} a_j b_k \mathbf{P}(A_j \cap B_k) \\
&= \sum_{j=1}^{m} \sum_{k=1}^{n} a_j b_k \mathbf{P}(A_j)\mathbf{P}(B_k) \\
&= \mathbf{E}\xi \cdot \mathbf{E}\eta.
\end{aligned}$$

2) 设 ξ, η 是独立的非负 r.v. , 令

$$\xi_n(\omega)=\begin{cases}\dfrac{j}{2^n}, & \omega\in A_j=\{\dfrac{j}{2^n}\leqslant\xi<\dfrac{j+1}{2^n}\},\\ & \qquad j=0,1,\cdots,n2^n-1,\\ n, & \omega\in A_{n2^n}=\{\xi\geqslant n\},\end{cases}$$

$$\eta_n(\omega)=\begin{cases}\dfrac{k}{2^n}, & \omega\in B_k=\{\dfrac{k}{2^n}\leqslant\xi<\dfrac{k+1}{2^n}\},\\ & \qquad k=0,1,\cdots,n2^n-1,\\ n, & \omega\in B_{n2^n}=\{\eta\geqslant n\}.\end{cases}$$

因为 ξ 与 η 独立，所以由定理 4 知事件类 $\{A_0, A_1, \cdots, A_{n2^n}\}$ 与 $\{B_0, B_1, \cdots, B_{n2^n}\}$ 独立. 注意到 ξ_n, η_n 是离散型 r.v. , 利用独立类扩张定理知 ξ_n, η_n 独立，此外，显然有 $0\leqslant\xi_n\uparrow\xi$, $0\leqslant\eta_n\uparrow\eta$, 因而 $0\leqslant\xi_n\eta_n\uparrow\xi\eta$, 利用 1) 及单调收敛定理有

$$\mathbf{E}\xi\eta=\lim_{n\to\infty}\mathbf{E}\xi_n\eta_n=\lim_{n\to\infty}\mathbf{E}\xi_n\mathbf{E}\eta_n=\mathbf{E}\xi\cdot\mathbf{E}\eta.$$

显然，如果 $\mathbf{E}\xi, \mathbf{E}\eta$ 有限，则 $\mathbf{E}\xi\eta$ 有限.

3) 设 ξ, η 是两个独立的实 r.v. 且数学期望有限. 由于 ξ, η 独立，根据独立类扩张定理，$(\xi^+, \xi^-, |\xi|)$ 与 $(\eta^+, \eta^-, |\eta|)$ 独立. 同时由于 $\mathbf{E}\xi, \mathbf{E}\eta$ 有限，因而 $|\xi|, |\eta|$ 可积，由 2) 知 $\mathbf{E}|\xi\eta|$ 有限. 现在往证 $\mathbf{E}\xi\eta=\mathbf{E}\xi\cdot\mathbf{E}\eta$. 注意到 (ξ^+, ξ^-) 与 (η^+, η^-) 的独立性，利用 2) 及可加性有

$$\begin{aligned}\mathbf{E}\xi\eta&=\mathbf{E}(\xi^+-\xi^-)(\eta^+-\eta^-)\\&=\mathbf{E}\xi^+\eta^+-\mathbf{E}\xi^+\eta^--\mathbf{E}\xi^-\eta^++\mathbf{E}\xi^-\eta^-\\&=\mathbf{E}\xi^+\mathbf{E}\eta^+-\mathbf{E}\xi^+\mathbf{E}\eta^--\mathbf{E}\xi^-\mathbf{E}\eta^++\mathbf{E}\xi^-\mathbf{E}\eta^-\\&=(\mathbf{E}\xi^+-\mathbf{E}\xi^-)(\mathbf{E}\eta^+-\mathbf{E}\eta^-)\\&=\mathbf{E}\xi\cdot\mathbf{E}\eta.\end{aligned}$$

4) 最后设 ξ, η 是两个独立的复 r.v. , 已知它们分别有有限的数学期望，仿 3) 可推出 $|\xi\eta|$ 可积，设 $\xi=\xi_1+i\xi_2$, $\eta=\eta_1+i\eta_2$,

其中 $\xi_1, \xi_2, \eta_1, \eta_2$ 是实 r.v.. 根据复 r.v. 独立性的定义，(ξ_1,ξ_2), (η_1,η_2) 独立，且它们的数学期望都有限，于是仿 3) 得到

$$\begin{aligned}
\mathbf{E}\xi\eta &= \mathbf{E}(\xi_1 + i\xi_2)(\eta_1 + i\eta_2) \\
&= \mathbf{E}\xi_1\eta_1 + i\mathbf{E}\xi_1\eta_2 + i\mathbf{E}\xi_2\eta_1 - \mathbf{E}\xi_2\eta_2 \\
&= \mathbf{E}\xi_1\mathbf{E}\eta_1 + i\mathbf{E}\xi_1\mathbf{E}\eta_2 + i\mathbf{E}\xi_2\mathbf{E}\eta_1 - \mathbf{E}\xi_2\mathbf{E}\eta_2 \\
&= (\mathbf{E}\xi_1 + i\mathbf{E}\xi_2)(\mathbf{E}\eta_1 + i\mathbf{E}\eta_2) \\
&= \mathbf{E}\xi \cdot \mathbf{E}\eta. \quad \square
\end{aligned}$$

6. 定理 设 $X_1, \cdots, X_n$ 为独立的 (实或复) r.v.，且 $\mathbf{E}|X_k|^2$ 有限，$k = 1, 2, \cdots, n$, 称

$$\mathbf{D}X_k := \mathbf{E}|X_k - \mathbf{E}X_k|^2$$

为 X_k 的 **方差**，则

$$\mathbf{D}X_k = \mathbf{E}|X_k|^2 - |\mathbf{E}X_K|^2, \tag{1}$$

$$\mathbf{D}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathbf{D}X_k. \tag{2}$$

证明 由 Schwarz 不等式知 $X_1, \cdots, X_n$ 有有限的数学期望，由积分性质易证第一个结论. 利用引理 1, $X_1 + \cdots + X_n$ 有有限的数学期望且

$$\mathbf{E}(X_1 + \cdots + X_n) = \mathbf{E}X_1 + \cdots + \mathbf{E}X_n.$$

由此，有

$$\begin{aligned}
&|(X_1 + \cdots + X_n) - \mathbf{E}(X_1 + \cdots + X_n)|^2 \\
&= \left[(X_1 - \mathbf{E}X_1) + \cdots + (X_n - \mathbf{E}X_n)\right] \\
&\quad \cdot \left[\overline{(X_1 - \mathbf{E}X_1)} + \cdots + \overline{(X_n - \mathbf{E}X_n)}\right] \\
&= |X_1 - \mathbf{E}X_1|^2 + \cdots + |X_n - \mathbf{E}X_n|^2 \\
&\quad + \sum_{i \neq j} (X_1 - \mathbf{E}X_i)\overline{(X_1 - \mathbf{E}X_j)}.
\end{aligned}$$

由于 X_i 与 $X_j(i \neq j)$ 独立，因而 $X_i - \mathbf{E}X_i$ 与 $\overline{X_j - \mathbf{E}X_j}$ 独立且有有限的数学期望，于是当 $i \neq j$ 时，

$$\mathbf{E}(X_i - \mathbf{E}X_i)\overline{(X_j - \mathbf{E}X_j)} = \mathbf{E}(X_i - \mathbf{E}X_i)\mathbf{E}\overline{(X_j - \mathbf{E}X_j)} = 0.$$

从而得到

$$\begin{aligned}&\mathbf{E}|(X_1 + \cdots + X_n) - \mathbf{E}(X_1 + \cdots + X_n)|^2 \\ &\quad = \mathbf{E}|X_1 - \mathbf{E}X_1|^2 + \cdots + \mathbf{E}|X_n - \mathbf{E}X_n|^2. \quad \square\end{aligned}$$

7. 定义 设 X 为实 r.v.，则称 $f_X(t) := \mathbf{E}e^{itX}$, $t \in \mathbb{R}$, 为 X 的 **特征函数**. 若 $X = (X_1, \cdots, X_n)$ 为 n 维实 r.v.，则称 $f_X(t) = \mathbf{E}e^{i(t,X)}$; $t \in \mathbb{R}^n$. 为 X **的特征函数**，其中 $t = (t_1, \cdots, t_n)$, $(t, x) = \sum_{k=1}^n t_k x_k$ (即 t, x 的内积).

因为对给定的 $t \in \mathbb{R}^n$, $e^{i(t,x)}$, $x \in \mathbb{R}^n$, 是 x 的连续函数，因而是 Borel 函数且有界，所以 $f_X(t)$ 是存在的.

8. 引理 若 $X^{(1)}, \cdots, X^{(m)}$ 是独立的 n 维实 r.v.，则对一切 $t^{(k)} \in \mathbb{R}^n$, $k = 1, 2, \cdots, m$, $X = (X^{(1)}, \cdots, X^{(m)})$ 的特征函数

$$(1) \qquad f_X(t) = \prod_{k=1}^m f_{X^{(k)}}(t^{(k)}), \qquad t = (t^{(1)}, \cdots, t^{(m)}) \in \mathbb{R}^{nm},$$

证明 由定理 4 知 $e^{i(t^{(1)}, x^{(1)})}, \cdots, e^{i(t^{(m)}, x^{(m)})}$ 是独立复 r.v.，所以由特征函数的定义及定理 5 知 (1) 式成立. $\square$

9. 定义 设 f 是 $(\Omega, \mathcal{F})$ 到 $(E, \mathcal{E})$ 上的可测映射，μ 是 $\mathcal{F}$ 上的测度，定义

$$(1) \qquad \mu_f(B) := \mu(f^{-1}(B)), \qquad \forall\, B \in \mathcal{E}.$$

则由定理 4.1.7 的证明方法知 μ_f 是 $\mathcal{E}$ 上的测度. (1) 有时记作

$$(2) \qquad \mu_f = \mu \circ f^{-1},$$

称为 μ 在 $(E, \mathcal{E})$ 上 **由 f 导出的测度**.

若 $(E,\mathcal{E})=(\overline{\mathbb{R}}^n,\overline{\mathcal{B}}^n)$, 则 $\mu_f=\mu\circ f^{-1}$ 称为 n 维可测函数 f 的分布.

若 μ 是概率, $(E,\mathcal{E})=(\mathbb{R}^n,\mathcal{B}^n)$, 则 $f=(f_1,\cdots,f_n)$ 是 n 维实 r.v. , $\mu_f=\mu\circ f^{-1}$ 就是 n 维 r.v. f 的 **概率分布**.

10. 定理 (积分变换定理) 设 f 是 $(\Omega,\mathcal{F},\mu)$ 到 $(E,\mathcal{E})$ 上的可测映射, g 是 $(E,\mathcal{E})$ 上的可测函数, 则

$$\int_{f^{-1}(B)}(g\circ f)\,d\mu=\int_B g\,d\mu_f,\qquad \forall B\in\mathcal{E}. \tag{1}$$

上式的意义是: 如等式的一边存在, 则另一边存在, 而且两者相等.

证明 首先令 $g=I_F$, $F\in\mathcal{E}$, 则

$$\begin{aligned}\int_B I_F\,d\mu_f&=\mu_f(B\cap F)=\mu(f^{-1}(B\cap F))\\&=\mu(f^{-1}(B)\cap f^{-1}(F))=\int_{f^{-1}(B)}I_F(f(\omega))\,\mu(d\omega)\\&=\int_{f^{-1}(B)}(I_F\circ f)\,d\mu.\end{aligned}$$

即 (1) 对 I_F 成立, 因而由积分的线性性质知 (1) 式当 g 为非负 $\mathcal{E}-$ 简单函数时成立. 若 g 为非负 $\mathcal{E}$ 可测函数, 则存在 $\{g_n:\ n\in\mathbb{N}\}$ 为非负 $\mathcal{E}-$ 简单函数序列, 且 $0\leqslant g_n\uparrow g$, 于是 $\{g_n\circ f:\ n\in\mathbb{N}\}$ 为非负 $\mathcal{F}-$ 简单函数列, 且 $0\leqslant g_n\circ f\uparrow g\circ f$, 因此由单调收敛定理的推论知 (1) 式对非负 $\mathcal{E}$ 可测函数成立.

若 g 为实 $\mathcal{E}$ 可测函数, 则

$$\int_{f^{-1}(B)}g^+\circ f\,d\mu=\int_B g^+\,d\mu_f,\qquad \forall B\in\mathcal{E}, \tag{2}$$

$$\int_{f^{-1}(B)}g^-\circ f\,d\mu=\int_B g^-\,d\mu_f,\qquad \forall B\in\mathcal{E}, \tag{3}$$

而 $g^{\pm}\circ f=(g\circ f)^{\pm}$. 故 $\int_B g\,d\mu_f$ 存在等价于 (2)、(3) 的右边有一有限，因而也等价于 (2)、(3) 的左边有一有限，即 $\int_{f^{-1}(B)} g\circ f\,d\mu$ 存在. 再由 (2)、(3) 及积分定义知 (1) 对实值 $\mathcal{E}$ 可测函数成立. 对于复值 $\mathcal{E}$ 可测函数可对其实部与虚部分别应用实值 $\mathcal{E}$ 可测函数的结果即得 (1) 式成立. □

11. 定义 设 μ 为 $(\mathbb{R}^n,\mathcal{B}^n)$ (或 $(\mathbb{R}^n,\mathcal{B}^n_\mu)$, $\mathcal{B}^n_\mu$ 是 $\mathcal{B}^n$ 关于 μ 的完全化) 上的测度, 且在有界集上的值有限, 称 μ 为 L-S(Lebesgue-tieltjes) 测度； $F(x)$, $x\in\mathbb{R}^n$ 是与 μ 对应的分布函数 (即定义在 $\mathbb{R}^n$ 上的有限值函数, 且满足： (1) $\Delta_{b,a}F\geqslant 0$, a、$b\in\mathbb{R}^n$, $a\leqslant b$; (2) $F(x_1,\cdots,x_n)$ 对每一坐标变量都是右连续的； (3) $\mu((a,b])=\Delta_{b,a}F$); f 是 $\mathcal{B}^n$ (或 $\mathcal{B}^n_\mu$) 可测函数, 则称 f 对 μ (或 F) 的积分为 L–S 积分, 记作

$$\int_{\mathbb{R}^n} f\,d\mu=\int\cdots\int_{\mathbb{R}^n} f(x_1,\cdots,x_n)\,\mu(dx_1,\cdots,dx_n)$$

$$\text{或}=\int_{-\infty}^{\infty}\cdots\int f(x_1,\cdots,x_n)\,dF(x_1,\cdots,x_n)$$

在 $B\in\mathcal{B}^n$ 上对 μ 的积分记作

$$\begin{aligned}\int_B f\,d\mu&=\int_B f(x)\,\mu(dx)\\&=\int\cdots\int_B f(x_1,\cdots,x_n)\,dF(x_1,\cdots,x_n).\end{aligned}$$

在区间 $(a,b]$ 上对 μ 的积分记作

$$\int_{(a,b]} f\,d\mu=\int_{a_1}^{b_1}\cdots\int_{a_n}^{b_n} f(x_1,\cdots,x_n)\,dF(x_1,\cdots,x_n),$$

其中 $a=(a_1,\cdots,a_n)$, $b=(b_1,\cdots,b_n)$.

若 μ 为 n 维 L- 测度 λ, 则分别记为:

$$\int_{\mathbb{R}^n} f\,d\lambda = \int_{-\infty}^{\infty}\cdots\int f(x_1,\cdots,x_n)\,dx_1\cdots dx_n,$$

$$\int_B f\,d\lambda = \underset{B}{\int\cdots\int} f(x_1,\cdots,x_n)\,dx_1\cdots dx_n,$$

$$\int_{(a,b]} f\,d\lambda = \int_{a_1}^{b_1}\cdots\int_{a_n}^{b_n} f(x_1,\cdots,x_n)\,dx_1\cdots dx_n.$$

下面我们来介绍积分变换定理对计算 r.v. 的期望及 r.v. 函数的期望的应用.

12. 定理 设 $X=(X_1,\cdots,X_n)$ 是 $(\Omega,\mathcal{F},\mathbf{P})$ 上的 n 维实 r.v., F 是它的概率分布函数, 则

1) 对一切 $B\in\mathcal{B}^n$

$$\mathbf{P}(X\in B) = \underset{B}{\int\cdots\int} dF(x_1,\cdots,x_n) = \int_B d\mathbf{P}_X. \tag{1}$$

2) 若 $g_k:\mathbb{R}^n\mapsto\mathbb{R}$ 为 $\mathcal{B}^n$ 可测函数, $Y_k := g_k(X) = g_k(X_1,\cdots,X_n)$, $k=1,2,\cdots,m$, 则 $Y=(Y_1,\cdots,Y_m)$ 的分布函数

$$\begin{aligned} F_Y(y_1,\cdots,y_m) &= \int_{G_y} d\mathbf{P}_X \\ &= \underset{G_y}{\int\cdots\int} dF(x_1,\cdots,x_n), \quad y=(y_1,\cdots,y_m)\in\mathbb{R}^m \end{aligned} \tag{2}$$

其中

$$G_y := \big\{(x_1,\cdots,x_n):\, g_k(x_1,\cdots,x_n)\leqslant y_k,\quad k=1,2,\cdots,m\big\}.$$

证明 (可由积分变换定理得到, 但下面的叙述更直观.)

因为 $\mathbf{P}(X\in B)=\mathbf{P}(X^{-1}(B))=\mathbf{P}_X(B)$, 由此即得 1). 只需

注意

$$\begin{aligned}F_Y(y_1,\cdots,y_m)&=\mathbf{P}(\{Y_1\leqslant y_1,\cdots,Y_m\leqslant y_m\})\\&=\mathbf{P}(\{g_k(X_1,\cdots,X_n)\leqslant y_k,\quad k=1,2,\cdots,m\})\\&=\mathbf{P}(X\in G_y),\end{aligned}$$

即得 2). □

下面是 r.v. 的函数的数学期望的 L–S 积分表示.

13. 定理 设 $X=(X_1,\cdots,X_n)$ 是 $(\Omega,\mathcal{F},\mathbf{P})$ 上的 n 维实 r.v., 其分布函数为 F_X, 又 g 为 $\mathcal{B}^n$ (实或复) 可测函数, 则

$$\begin{aligned}\mathbf{E}g(X)&=\int_\Omega g\circ X\,d\mathbf{P}=\int_{\mathbb{R}^n}g\,d\mathbf{P}_X\\&=\int\cdots\int_{\mathbb{R}^n} g(x_1,\cdots,x_n)\,dF(x_1,\cdots,x_n).\end{aligned}\tag{1}$$

对于 X 的特征函数 $f_X(t)=\mathbf{E}[e^{i\sum_{k=1}^n t_kX_k}]$, 有

$$\begin{aligned}f_X(t)&=\int\cdots\int_{\mathbb{R}^n} e^{i\sum_{k=1}^n t_kx_k}\,dF_X(x_1,\cdots,x_n)\\&=\int_{\mathbb{R}^n}e^{i(t,x)}\mathbf{P}_X(dx),\quad t=(t_1,\cdots,t_n)\in\mathbb{R}^n.\end{aligned}\tag{2}$$

证明 由积分变换定理及定义 11 立得.

14. 注 1) 对于一个随机变量来说, 它的 L–S 积分可以有很多种, 例如拿定理 13 中 X 的第一个分量 X_1 来说, 如果令

$$g(x_1,\cdots,x_n):=x_1,$$

则由 (13.1) 得

$$\mathbf{E}X_1=\int_{-\infty}^{\infty}\cdots\int x_1\,dF_X(x_1,\cdots,x_n).\tag{1}$$

如果令定理 13 的 $n=1$, $X=X_1$, $g(x)=x$, $x\in\mathbb{R}$, 则由 (13.1) 即得

$$\mathbf{E}X_1=\int_{-\infty}^{\infty}x\,dF_{X_1}(x). \tag{2}$$

这样就得到 $\mathbf{E}X_1$ 的两个 L–S 积分表达式，它们是一致的，这个一致性也可以用 L–S 的方法证明. 也可从概率的直观来说明如下：实际上，$F_{X_1}(x_1)$ 是 $F_X(x_1,\cdots,x_n)$ 的关于 X_1 的边缘分布. 在 (1) 中对 $x_2,\cdots,x_n$ 的 $n-1$ 重积分与 x_1 无关，实际上这 $n-1$ 重积分就是将 $dF_X(x_1,\cdots,x_n)$ 化为对 x_1 的一重积分，即化为 $dF_{X_1}(x_1)$, 所以 (1), (2) 应该是一致的.

注 2) 应用定理 13 也可以证明期望的有关性质，例如定理 1 可以如下证明 (令 $X=(X_1,\cdots,X_n)$): 由 (13.1)

$$\begin{aligned}\mathbf{E}(c_1X_1+\cdots+c_nX_n)&=\int\underset{\mathbb{R}^n}{\cdots}\int(c_1x_1+\cdots+c_nx_n)\,dF_X(x_1,\cdots,x_n)\\&=\sum_{k=1}^{n}c_k\int\underset{\mathbb{R}^n}{\cdots}\int x_k\,dF_X(x_1,\cdots,x_n)\\&=\sum_{k=1}^{n}c_k\mathbf{E}X_k.\end{aligned}$$

注 3) 定理 12, 13 是对一般 r.v. 而言的，对于离散 r.v. 来说，公式 (12.1), (12.2), (13.1), (13.2) 有更简单的形式，由于方法是一样的，所以只考虑一维情形已经足够. 例如设 $I\subset\mathbb{N}$,

$$X=\sum_{k\in I}a_kI_{\{X=a_k\}},$$

则其概率分布可以用

$$\mathbf{P}(\{X=a_k\})=\mathbf{P}_X(\{a_k\})=p_k,\qquad k\in I,$$

表示，而 (12.1)、(12.2), (13.1), (13.2) 可以分别写成

$$\mathbf{P}(X \in B) = \sum_{k:\, a_k \in B} \mathbf{P}_X(\{a_k\}) = \sum_{k:\, a_k \in B} p_k;$$

$$\mathbf{P}(g(X) \leqslant y) = \sum_{k:\, g(a_k) \leqslant y} \mathbf{P}_X(\{a_k\}) = \sum_{k:\, g(a_k) \leqslant y} p_k;$$

$$\mathbf{E}g(X) = \sum_{k \in I} g(a_k)\mathbf{P}_X(\{a_k\}) = \sum_{k \in I} g(a_k) p_k;$$

$$f_X(t) = \sum_{k \in I} e^{ita_k}\mathbf{P}_X(\{a_k\}) = \sum_{k \in I} e^{ita_k} p_k.$$

对于连续型 r.v.，则可以将上述各式表示成关于分布密度的积分. 为此我们先证明下述定理.

15. 定理 设 μ 是 $(\Omega, \mathcal{F})$ 上的测度， p 是非负 $\mathcal{F}-$ 可测函数，定义

$$\nu(A) := \int_A p(\omega)\,\mu(d\omega), \quad A \in \mathcal{F}, \tag{1}$$

则 ν 是 $\mathcal{F}$ 上的测度，且 $\int g(\omega)\nu(d\omega)$ 存在当且仅当 $\int g(\omega)p(\omega)\mu(d\omega)$ 存在， 当积分存在时，

$$\int_A g(\omega)\nu(d\omega) = \int_A g(\omega)p(\omega)\mu(d\omega), \quad \forall\, A \in \mathcal{F}. \tag{2}$$

证明 1) 显然 ν 是 $\mathcal{F}$ 上的非负集函数，往证其 $\sigma-$ 可加性，设 $A = \bigcup_{n=1}^{\infty} A_n$; $A_n \in \mathcal{F}$、$n \in \mathbb{N}$, 两两不交.

$$\begin{aligned}
\nu(A) &= \int_A p(\omega)\,\mu(d\omega) = \int I_A(\omega)p(\omega)\,\mu(d\omega)\\
&= \int I_{\bigcup_{n=1}^{\infty} A_n(\omega)} p(\omega)\,\mu(d\omega)\\
&= \lim_{N\to\infty} \int I_{\bigcup_{n=1}^{N} A_n}(\omega)p(\omega)\,\mu(d\omega) \qquad \text{(单调收敛定理)}\\
&= \lim_{N\to\infty} \int \sum_{n=1}^{N} I_{A_n}(\omega)p(\omega)\,\mu(d\omega) \quad (A_n,\quad n \in \mathbb{N}\text{两两不交})
\end{aligned}$$

$$= \lim_{N\to\infty} \sum_{n=1}^{N} \int I_{A_n}(\omega)p(\omega)\,\mu(d\omega)$$

$$= \lim_{N\to\infty} \sum_{n=1}^{N} \nu(A_n) = \sum_{n=1}^{\infty} \nu(A_n).$$

2) 当 $g = I_B$, $B \in \mathcal{F}$ 时, 对一切 $A \in \mathcal{F}$

$$\int_A g\,d\nu = \int I_{A\cap B}\,d\nu = \nu(A\cap B)$$

$$= \int_{A\cap B} p(\omega)\,\mu(d\omega) = \int_A g(\omega)p(\omega)\,\mu(d\omega).$$

故 (2) 对示性函数成立, 从而由积分线性性质及单调收敛定理知 (2) 式对一切非负可测函数 g 成立 (此时积分的存在性不成问题).

设 g 为实可测函数, 则 $(gp)^\pm = g^\pm p$. 由已证结论知

$$\int_A g^\pm d\nu = \int_A g^\pm p d\mu, \quad \forall A \in \mathcal{F}. \tag{3}$$

所以 $\int g\,d\nu$ 存在等价于 $\int g^+ p d\nu = \int (gp)^- d\mu$, $\int g^-\,d\nu = \int (gp)^- d\mu$ 有一有限, 亦即等价于 $\int gpd\mu$ 存在, 而且由 (3) 知 (2) 成立.

对复可测函数 g, 只要分别对其实、虚部应用 (2) 式, 然后综合成 g 的积分即得所需结论. □

16. 定理 设 $X = (X_1, \cdots, X_n)$ 为具有分布密度 $p(x) = p(x_1, \cdots, x_n)$ 的连续型 r.v..

1) $g(x) = g(x_1, \cdots, x_n)$ 是 n 元 Borel 可测函数, 则 $\mathbf{E}g(X)$ 存在的充要条件是 $g(x)p(x)$ 在 $\mathbb{R}^n$ 上对 L- 测度积分存在. 当 $\mathbf{E}g(X)$ 存在时,

$$\mathbf{E}g(X) = \int_{\mathbb{R}^n} g(x)p(x)\,dx \tag{1}$$

$$= \int_{-\infty}^{\infty} \cdots \int g(x_1, \cdots, x_n)p(x_1, \cdots, x_n)\,dx_1 \cdots dx_n,$$

特别

$$\mathbf{P}(X \in B) = \int \cdots \int_B p(x_1, \cdots, x_n)\, dx_1 \cdots dx_n. \tag{2}$$

2) 若 $g_k: \mathbb{R}^n \to \mathbb{R}$ 为 $\mathcal{B}^n$– 可测函数, $Y_k := g_k(X)$, $k = 1, 2, \cdots, m$, 则 $Y = (Y_1, \cdots, Y_m)$ 的分布函数

$$F_Y(y_1, \cdots, y_m) = \int \cdots \int_{G_y} p(x_1, \cdots, x_n)\, dx_1 \cdots dx_n, \tag{3}$$

其中 G_y 由定理 12(b) 定义.

$$\begin{aligned} \mathbf{P}_Y(B) &= \mathbf{P}(\{Y \in B\}) \\ &= \int \cdots \int_{\{x \in \mathbb{R}^n : g(x) \in B\}} p(x_1, \cdots, x_n)\, dx_1 \cdots dx_n, \quad B \in \mathcal{B}^m, \end{aligned} \tag{4}$$

其中 $g(x) := (g_1(x), \cdots, g_m(x))$.

证明 在定理 15 中令 $\Omega = \mathbb{R}^n$, $\mathcal{F} = \mathcal{B}^n$, μ 为 $\mathbb{R}^n$ 上的 L 测度, $\nu = \mathbf{P}_X$, 则由定理 13 及 15 即得 (1), 再分别令 $g = I_B$、I_{G_y}、$I_{\{x \in \mathbb{R}^n, g(x) \in B\}}$, 即得 (2), (3), (4). □

17. 关于 Lebesgue 积分与 Riemann 积分间的关系的注: 在定理 16 中的积分都是 Lebesgue 积分. 要进行实际计算, 通常是对 Riemann 积分进行的. 因此需要讨论 L– 积分与 R– 积分之间的关系. 为了简单起见, 我们先考察 $\mathbb{R}$ 中有限区间 $[a, b]$ 上有界可测函数 f 的 L– 积分与 R– 积分的关系.

由于 f 是 $([a, b], \mathcal{B}[a, b])$ 上的有界可测函数, 所以它的 L– 积分 (暂记作 $(L)\int_{[a,b]} f(x)\, dx$) 总是存在的. 按照 R– 积分的定义, f 的 R– 积分 (记作 $(R)\int_a^b f(x)\, dx$) 存在 (也称为 R 可积) 的充要条件是对于 $[a, b]$ 的任一分划

$$T: \ a = x_0 < x_1 < \cdots < x_n = b,$$

记

$$\ell(T) = \max_{1\leqslant k\leqslant n}(x_k - x_{k-1}),$$
$$m_k = \min_{x_{k-1}\leqslant x\leqslant x_k} f(x),$$
$$M_k = \max_{x_{k-1}\leqslant x\leqslant x_k} f(x).$$

有

$$\lim_{\ell(T)\to 0}\sum_{k=1}^n m_k(x_k - x_{k-1}) = \lim_{\ell(T)\to 0}\sum_{k=1}^n M_k(x_k - x_{k-1}).$$

而在 R 可积的条件下，上述共同极限即为 $(R)\int_a^b f(x)\,dx$.

另一方面，根据上述分划 T，可以定义两个 $\mathcal{B}[a,b]$ 简单函数

$$\underline{f_T} := m_1 I_{[x_0,x_1]} + \sum_{k=2}^n m_k I_{(x_{k-1},x_k]},$$

$$\overline{f_T} := M_1 I_{[x_0,x_1]} + \sum_{k=2}^n M_k I_{(x_{k-1},x_k]},$$

对此分划 T,

$$\underline{f_T}(x) \leqslant f(x) \leqslant \overline{f_T}(x), \qquad \forall\, x\in[a,b].$$

用 λ 表示 L- 测度，则

$$\sum_{k=1}^n m_k(x_k - x_{k-1}) = (L)\int_{[a,b]}\underline{f_T}\,d\lambda \leqslant (L)\int_{[a,b]} f\,d\lambda$$
$$\leqslant (L)\int_{[a,b]}\overline{f_T}\,d\lambda = \sum_{k=1}^n M_k(x_k - x_{k-1}).$$

令 $\ell(T)\to 0$ 即得

$$(R)\int_a^b f(x)\,dx = (L)\int_{[a,b]} f\,d\lambda.$$

按照相同的方法，不难证明 (细节参见 [YWL] III.5.2, pp.203–210) 下列命题.

1) 设 G 为 $\mathbb{R}^n$ 中的一有界可求积区域， $f: G \mapsto \mathbb{R}$ (或 $\mathbb{C}$) 是 $\mathcal{B}^n \cap G$ 有界可测函数． 若 f 在 G 上 R- 可积， 则 f 在 G 上 L- 可积， 且 f 在 G 上的 L- 积分与 R- 积分相等， 即

$$(R)\underset{G}{\int\cdots\int} f(x_1,\cdots,x_n)dx_1\cdots dx_n = (L)\int_G f\,d\lambda.$$

其中 λ 是 L- 测度．

2) 设 $G \subset \mathbb{R}^n$ 为无界区域，它与 $\mathbb{R}^n$ 的任一有界可求积区域的交可求积， $f: G \mapsto \mathbb{R}$ (或 $\mathbb{C}$) 是 $\mathcal{B}^n(G)$ 可测函数，若 f 在 G 上 R- 绝对可积 (即在 G 上， $|f|$ 可积)，且在 G 的每一有界可求积子区域上 R- 可积，则 f 在 G 上 L- 可积，且 f 在 G 上的 L- 积分与 R- 积分相等． (关于无界函数的广义积分有类似的性质．)

有了这两条命题，如果 R- 积分可以进行实际计算，就可以将 L- 积分转化成 R- 积分来计算．

在 R- 积分的计算中，变量替换公式是一个重要工具，借助于 L- 积分与 R- 积分的关系也可得到 L - 积分的变量替换公式如下：

3) 设 $(\mathbb{R}^n_x, \mathcal{B}^n_x, \lambda_x)$ 和 $(\mathbb{R}^n_t, \mathcal{B}^n_t, \lambda_t)$ 是 L- 测度空间， $D_x \in \mathcal{B}^n_x$, $D_t \in \mathcal{B}^n_t$, 在 D_x 与 D_t 之间存在一个一一映射： $x = x(t) \in D_x$, 对一切 $t \in D_t$, 即

$$x = (x_1,\cdots,x_n) = (x_1(t_1,\cdots,t_n),\cdots,x_n(t_1,\cdots,t_n))$$

$x_k = x_k(t_1,\cdots,t_n)$, $k = 1,2,\cdots,n$, 在 D_t 上存在连续偏导数且

$$J(t) = \frac{\partial(x_1,\cdots,x_n)}{\partial(t_1,\cdots,t_n)} \neq 0, \qquad \forall t = (t_1,\cdots,t_n) \in \bar{D}_t.$$

再设 D_x 与 D_t 是有界可求积区域或为无界区域但与 $\mathbb{R}^n$ 的任一有界可求积区域之交可求积． $f: D_x \mapsto \mathbb{R}$ 为 $D_x \cap \mathcal{B}^n_x$ 可测函数，

则 $\int_{D_x} f(x)\,\lambda_x(dx)$ 存在当且仅当 $\int_{D_t} f(x(t))|J(t)|\,\lambda_t(dt)$ 存在，且在此情况下，

$$\int_{D_x} f(x)\,\lambda_x(dx) = \int_{D_t} f(x(t))|J(t)|\,\lambda_t(dt).$$

公式两边，可以理解为 L- 积分．由命题 (2)，当 R- 积分存在时，又可以理解为 R- 积分．

由上述命题 1)–3)．可以证明下列的

18．定理　设 $X=(X_1,\cdots,X_n)$ 是 n 维连续型 r.v.，其分布密度为 $p_X(x)\doteq p_X(x_1,\cdots,x_n)$．再设 $\mathbb{R}^n$ 上的 n 元可测函数 $y=y(x)\in\mathbb{R}^n_y$、$x\in\mathbb{R}^n_x$．(即 $y=(y_1,\cdots,y_n)$，$y_k=y_k(x)=y_k(x_1,\cdots,x_n)$，$x=(x_1,\cdots,x_n)\in\mathbb{R}^n_x$) 满足下述条件：除 $\mathbb{R}^n_x$ 的 L- 零测集 N 外，存在 $\{D^x_k:\ k\in I\}\subset\mathcal{B}^n_x$，$I\subset\mathbb{N}$，两两不交，满足 (17.3) 的要求，且 $\bigcup_{k\in I}D^x_k=\mathbb{R}^n_x\setminus N$，对一切 $k\in I$，映射 $y:\ D^x_k\mapsto D^y_k$ 是一一的，D^y_k，$k\in I$ 满足 (17.3) 的要求但不一定不相交，在每个 D^x_k 上，y 的逆映射 $x^{(k)}=x^{(k)}(y)$ 的分量是 D^y_k 上具连续偏导数的函数，且函数行列式

$$J_k(y)=\frac{\partial(x_1^{(k)},\cdots,x_n^{(k)})}{\partial(y_1,\cdots,y_n)}\neq 0,\qquad \forall y\in D^y_k.$$

若 $Y=(Y_1,\cdots,Y_n)$，$Y_k=y_k(X)$，$k=1,2,\cdots,n$，则 Y 是连续型 r.v.，其分布密度为

$$p_Y(y_1,\cdots,y_n)=\sum_{k\in I}I_{D^y_k}(y)p_X(x^{(k)}(y))\cdot|J_k(y)|.$$

19．例　设 X 的分布密度为 $p_X(x)$，试求 $Y=\sin X$ 的分布密度，当 X 服从 $[-\frac{\pi}{2},\frac{\pi}{2}]$ 上的均匀分布时，Y 的分布密度又如何？

解： 由于 Y 只在 $[-1,1]$ 中取值，而 $y=\sin x$ 的逆 (反) 函数

有无穷个分支:

$$x^{(k)}(y) = k\pi + (-1)^k \sin^{-1} y, \qquad k \in \mathbb{Z}.$$

其中 $\sin^{-1}(y) \in (-\frac{\pi}{2}, \frac{\pi}{2})$, 当 $y \in (-1, 1)$. 于是 $D_k^x = k\pi + (-\frac{\pi}{2}, \frac{\pi}{2})$

$$\mathbb{R}_x^1 \setminus \left\{\frac{k}{2}\pi:\ k \in 2\mathbb{Z}-1\right\} = \bigcup_{k\in\mathbb{Z}} D_k^x.$$

$D_k^y = (-1, 1)$, $k \in \mathbb{Z}$. 即满足定理的要求. 而

$$\begin{aligned}\frac{dx^{(n)}(y)}{dy} &= (-1)^n \frac{d\sin^{-1} y}{dy} = (-1)^n \left(\frac{dy}{dx}\right)^{-1} \\ &= \frac{(-1)^n}{\cos x} = \frac{(-1)^n}{\sqrt{1-y^2}},\end{aligned}$$

故

$$p_Y(y) = \begin{cases} 0, & |y| > 1, \\ \frac{1}{\sqrt{1-y^2}} \sum\limits_{n=-\infty}^{\infty} p_X(n\pi + (-1)^n \sin^{-1} y), & |y| \leqslant 1. \end{cases}$$

当 X 在 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 上服从均匀分布时,

$$p_X(x) = \begin{cases} \frac{1}{\pi}, & |x| \leqslant \frac{\pi}{2} \\ 0, & |x| > \frac{\pi}{2} \end{cases}$$

则

$$p_Y(y) = \begin{cases} 0, & |y| > 1 \\ \frac{1}{\pi\sqrt{1-y^2}}, & |y| \leqslant 1. \end{cases}$$

(注: 也可以直接应用定理.)

习题

1. 设 ξ, η 为实 r.v., $\mathbf{E}\xi^2$, $\mathbf{E}\eta^2$ 有限, 试证 $\mathbf{D}(\xi+\eta) = \mathbf{D}\xi + \mathbf{D}\eta$ 的充要条件是 $\mathbf{E}(\xi\eta) = \mathbf{E}\xi \cdot \mathbf{E}\eta$.

2. 若 $X_1, \cdots, X_n$ 是 n 个实 r.v.，其分布函数分别为 $F_1(x), \cdots, F_n(x)$，η 是一 r.v.，其分布函数为 $\frac{1}{n}(F_1(x)+\cdots+F_n(x))$. 设 r 为一正数，证明

$$\mathbf{E}|\eta|^r = \frac{1}{n}\sum_{k=1}^{n}\mathbf{E}|X_n|^r.$$

3. 设 X 是一实 r.v.，它的分布函数是 $F(x)$，c 为一正数，令

$$X^c := \begin{cases} X, & \text{当}|X| < c, \\ c, & \text{当}X \geqslant c, \\ -c, & \text{当}X \leqslant -c. \end{cases}$$

试将 $\mathbf{E}X^c$、$\mathbf{D}X^c$ 用对于 F 的 L S 积分表出.

4. 设 $F(x)$ 是一分布函数，按定义

$$\int_a^b f(x)\,dF(x) = \int_{(a,b]} f(x)\,dF(x),$$

问它是否等于 $\int_{[a,b]} f(x)\,dF(x)$？在什么情况下它们不相等？

5. 设 X 是一实 r.v.，m 是一实数，它满足 $\mathbf{P}(\{X \geqslant m\}) \geqslant \frac{1}{2}$，$\mathbf{P}(\{X \leqslant m\}) \geqslant \frac{1}{2}$.（称满足上述条件的 m 为 X 的 **中数**）试证：

(1) 对一切 $a \in \mathbb{R}$，$\mathbf{E}|X - m| \leqslant \mathbf{E}|X - a|$；

(2) X 的中数，数学期望和方差之间有下述关系：

$$\mathbf{E}X - \sqrt{\mathbf{D}X} \leqslant m \leqslant \mathbf{E}X + \sqrt{\mathbf{D}X}$$

6. 设 $X = \sum_{k\in I} a_k I_{\{X=a_k\}}$，$a_k \in \mathbb{R}$，$I \subset \mathbb{N}$，试证：

$$\mathbf{P}(X \in B) = \sum_{k\in I,\, a_k\in B} \mathbf{P}_X(\{a_k\}),$$

$$\mathbf{E}g(X) = \sum_{k\in I} g(a_k)\mathbf{P}_X(\{a_k\}).$$

以下各题要经常用到下列结论：设 X_k 具有分布密度 $p_x(x_k)$，$X_1, \cdots, X_n$ 独立，则 $(X_1, \cdots, X_n)$ 具有分布密度 $p_1(x_1)\cdots p_n(x_n)$.

7. 设 X_1, X_2 是在 $[a,b]$ 上均匀分布的独立 r.v.，求 $Y = X_1 + X_2$ 的分布函数与分布密度 (可以由公式用数学分析计算，也可用几何求面积方法计算).

8. 设 X, Y 为独立 r.v.，X 在 $[0,1]$ 上均匀分布，Y 按二项分布律 $B(n,p)$ 分布，试证 $X+Y$ 是连续型随机变量，并求其分布密度.

9. 1) 设 $X = (X_1, X_2)$ 的分布密度为 $p_X(x_1, x_2)$,

$$Y_1 := \sqrt{X_1^2 + X_2^2}, \quad Y_2 := \frac{X_1}{X_2},$$

试求 $Y = (Y_1, Y_2)$ 的分布密度 $p_Y(y_1, y_2)$.

2) 设 $X = (X_1, X_2)$ 的分布密度 $p_X(x_1, x_2) := \frac{1}{2\pi\sigma^2} e^{-\frac{x_1^2 + x_2^2}{2\sigma^2}}$，试求 (1) 中定义的 $Y = (Y_1, Y_2)$ 的分布密度，并证明 Y_1, Y_2 独立.

10. 设 $X_1, \cdots, X_n$ 是具有相同分布密度 $p(x)$ 的独立 r.v.，且当 $x < 0$ 时 $p(x) = 0$，当 $x \geqslant 0$ 时 $p(x)$ 连续. 其次设 ξ_k^* 为 $X_1, \cdots, X_n$ 按不降顺序排列的第 k 个值，试证 $(\xi_1^*, \cdots, \xi_r^*)$, $1 \leqslant r \leqslant n$ 的分布密度为

$$p(y_1, \cdots, y_r) = \begin{cases} \frac{n!}{(n-r)!} p(y_1) \cdots p(y_r) \left(\int_{y_r}^{\infty} p(x)\, dx \right)^{n-r}, & 0 \leqslant y_1 \leqslant \cdots \leqslant y_r, \\ 0, & \text{其它}. \end{cases}$$

11. 设 $F(x)$ 和 $G(x)$ 是两个有界分布函数 (不一定是概率分布函数), $G(-\infty) = 0$, $f(x)$ 是连续函数，且 $0 < c_1 \leqslant f(x) < c_2 < \infty$, 对一切 $x \in \mathbb{R}$. 试应用定理 10 证明：若

$$F(x) = \int_{-\infty}^{x} \frac{1}{f(x)}\, dG(x), \qquad \forall x \in \mathbb{R},$$

则

$$G(x) = \int_{-\infty}^{x} f(x)\, dF(x), \qquad \forall\, x \in \mathbb{R}.$$

12. 设 $X_1, X_2, \cdots$ 是无穷个独立 r.v. (即其中任意有限个都独立), 且它们的分布都是以 $\lambda(>0)$ 为参数的指数分布, 即 $\mathbf{P}(X_k > t) = e^{-\lambda t}$, $t \geqslant 0$. 令

$$N(t) := \sup\left\{n \in \mathbb{N}: \sum_{k=1}^{n} X_k \leqslant t\right\}.$$

试证: 1) 若 $0 < s < t < \infty$, 则 $N(s)$ 与 $N(t) - N(s)$ 独立, 且分别服从参数为 λs 及 $\lambda(t-s)$ 的 Poisson 分布.

2) 对任何 m 及任何 $0 = t_0 < t_1 < \cdots < t_m < \infty$, $N(t_k) - N(t_{k-1})$, $k = 1, 2, \cdots, m$, 独立. 且 $N(t_k) - N(t_{k-1})$ 服从以 $\lambda(t_k - t_{k-1})$ 为参数的 Poisson 分布, $k = 1, 2, \cdots, m$.

13. 设 $X_1, X_2, \cdots, Y_1, Y_2, \cdots$ 是无穷个独立的非负 r.v., X_k 都服从以 $\lambda(>0)$ 为参数的指数分布. Y_k 服从集中在 $(0, \infty)$ 上的分布 μ, 令 $N(t)$ 为一 r.v., 它满足:

$$\{N(t) = n\} = \{X_1 + \cdots + X_n + Y_1 + \cdots + Y_n \leqslant t,$$
$$X_1 + \cdots + X_{n+1} + Y_1 + \cdots + Y_n > t\}$$

试证:

$$\mathbf{P}(\{N(t) = n\}) = \int_{(0,t]} \frac{[\lambda(t-x)]^n}{n!} \mu^{*n}(dx).$$

其中

$$\mu^{*n}((0, x]) := \mathbf{P}(\{Y_1 + \cdots + Y_n \leqslant x\}),$$

即 μ 的 n 重卷积.

14. 设 $X_1, \cdots, X_n$ 是独立 r.v., 且都服从 $(0,1)$ 上的均匀分布, 试应用广义加法公式 (或称逐步淘汰原则) 及公式

$$\int \cdots \int_{\sum_{k=1}^{n} x_k \leqslant x,\, x_i \geqslant 0} dx_1 \cdots dx_n = \frac{x^n}{n!},$$

证明：当 $x \in (0,n)$ 时，

$$\mathbf{P}(X_1 + \cdots + X_n \leqslant x) = \sum_{k=0}^{n-1} (-1)^k \binom{n}{k} \frac{[(x-k) \vee 0]^n}{n!}.$$

§4. 积分收敛定理

积分号下取极限是测度论中的一个基本问题，在 §1 中我们已经证明了单调收敛定理，它是积分收敛定理之一，已在 §2 中有所应用. 在这一节将再证明两个著名的定理并列举一些推论，它们在测度论，概率论以至分析数学中都有重要的作用.

首先将 §1 证明过的单调收敛定理加以简单推广.

1. 定理 (单调收敛定理) 给定 $(\Omega, \mathcal{F}, \mu)$，$g$ 为实可积函数，$f_n, n \in \mathbb{N}$，是实 $\mathcal{F}$ 可测函数，若 $g \leqslant f_n \uparrow f$，a.e.，则

$$\lim_{n\to\infty} \int f_n = \int \lim_{n\to\infty} f_n. \tag{1}$$

证明 由 g 可积知 g a.e. 有限，所以 $0 \leqslant f_n - g \uparrow f - g$，a.e. 成立，由定理 1.5 及积分的线性性质知

$$\begin{aligned}
\lim_{n\to\infty} \int f_n &= \lim_{n\to\infty} \left[\int (f_n - g) + \int g \right] \\
&= \int \lim_{n\to\infty} (f_n - g) + \int g \\
&= \int \left(\lim_{n\to\infty} f_n - g \right) + \int g \\
&= \int \lim_{n\to\infty} f_n..
\end{aligned}$$

2. 定理 (Fatou 引理) 给定 $(\Omega, \mathcal{F}, \mu)$，设 g, h 是实值可积函数，$\{f_n : n \in \mathbb{N}\}$ 是实可测函数列.

1) 对一切 $n \in \mathbb{N}$, $f_n \geqslant g$, a.e., 则

$$\int \varliminf_{n\to\infty} f_n \leqslant \varliminf_{n\to\infty} \int f_n.$$

2) 对一切 $n \in \mathbb{N}$, $f_n \leqslant h$, a.e., 则

$$\int \varlimsup_{n\to\infty} f_n \geqslant \varlimsup_{n\to\infty} \int f_n.$$

证明 首先由假设知对一切 $n \in \mathbb{N}$, $\int f_n$ 存在 (例如当 $f_n \geqslant g$ 时, $f_n^+ \geqslant g^+$, $f_n^- \leqslant g^-$, 因而 $\int f_n^- \leqslant \int g^- < \infty$, 故 $\int f_n$ 存在). 先证 1) 成立: 此时 $f_n \geqslant g$, 令

$$g_n := \inf_{k \geqslant n} f_k.$$

则 $g \leqslant g_n \uparrow \sup_n \inf_{k\geqslant n} f_k = \varliminf_{n\to\infty} f_n$, 故由定理 1 知

$$\lim_{n\to\infty} \int g_n = \int \lim_{n\to\infty} g_n = \int \varliminf_{n\to\infty} f_n.$$

而由 g_n 的定义知 $g_n \leqslant f_n$, 所以 $\int g_n \leqslant \int f_n$, $\lim_{n\to\infty} \int g_n \leqslant \varliminf_{n\to\infty} \int f_n$, 故 1) 成立.

其次由于 $\varlimsup_{n\to\infty} f_n = -\varliminf_{n\to\infty}(-f_n)$ 以及 $-h \leqslant -f_n$, $-h$ 可积, 所以由 1) 知

$$\begin{aligned}\varlimsup_{n\to\infty} \int f_n &= -\varliminf_{n\to\infty}\left(-\int f_n\right) = -\varliminf_{n\to\infty}\int(-f_n)\\ &\leqslant -\int \varliminf_{n\to\infty}(-f_n) = \int \varlimsup_{n\to\infty} f_n.\end{aligned}$$

故 2) 获证.

3. 定理 (控制收敛定理). 给定 $(\Omega, \mathcal{F}, \mu)$, g, h 为实可积函数.

1) 若 $f_n, n \in \mathbb{N}$ 为实可测函数序列, 当 $g \leqslant f_n \leqslant h$, a.e., 对一切 $n \in \mathbb{N}$, $f_n \xrightarrow{\text{a.e.}} f$ 时, 有 $\int f_n \to \int f$.

2) 若 $f_n, n \in \mathbb{N}$ 为实或复可测函数, 当 $|f_n| \leqslant g$, a.e., 对一切 $n \in \mathbb{N}$, $f_n \to f$, a.e. 时, 有 $\int |f_n - f| \to 0$, 因而 $\int f_n \to \int f$.

证明 由 1) 的假设知 Fatou 引理的 1), 2) 的假设都成立，而且有 $f=\underline{\lim}_{n\to\infty}f_n=\overline{\lim}_{n\to\infty}f_n$, a.e., 所以有

$$\overline{\lim_{n\to\infty}}\int f_n\leqslant\int\overline{\lim_{n\to\infty}}f_n=\int f=\int\underline{\lim_{n\to\infty}}f_n\leqslant\underline{\lim_{n\to\infty}}\int f_n,$$

即 1) 的结论成立.

其次由 2) 的假设知 $|f_n-f|\to 0$, a.e., $0\leqslant|f_n-f|\leqslant 2g$, a.e., $n\in\mathbb{N}$, 所以由 1) 得 $\int|f_n-f|\to 0$, 再由 $|\int f_n-\int f|\leqslant\int|f_n-f|$ 即得 $\int f_n\to\int f$

注 控制收敛定理还可以将条件 $f_n\to f$, a.e., 换成 $f_n\overset{\mu}{\to}f$. 这要用到有关函数列依测度收敛的概念，留待以后再讨论.

4. 推论 给定 $(\Omega,\mathcal{F},\mu)$, 设 f_n, $n\in\mathbb{N}$ 是 $\mathcal{F}-$ 可测函数列，若 f_n, $n\in\mathbb{N}$ 非负或 $\sum_{n=1}^{\infty}\int|f_n|\,d\mu<\infty$, 则 $\sum_{n=1}^{\infty}f_n$ 收敛，积分存在且

$$\int\sum_{n=1}^{\infty}f_n=\sum_{n=1}^{\infty}\int f_n. \tag{1}$$

注 条件可以减弱为 $\sum_{n=1}^{\infty}f_n$, a.e. 收敛且 $|\sum_{k=1}^{n}f_k|\leqslant g$, a.e., 对一切 $n\in\mathbb{N}$, g 可积.

证明 令 $g_n:=\sum_{k=1}^{n}|f_k|$ (当 f_n, $n\in\mathbb{N}$ 非负时 $g_n=\sum_{k=1}^{n}f_k$), 则 $0\leqslant g_n\uparrow g:=\sum_{n=1}^{\infty}|f_n|$, 由积分的线性性质及单调收敛定理，当 f_n, $n\in\mathbb{N}$ 非负时即得 (1) 式；当 $\sum_{n=1}^{\infty}\int|f_n|<\infty$ 时， $g=\sum_{n=1}^{\infty}|f_k|$ 可积，因而 a.e. 有限，故 $\sum_{n=1}^{\infty}f_n$, a.e. 收敛且对一切 $n\in\mathbb{N}$, $|\sum_{k=1}^{n}f_k|\leqslant g$. 由积分的线性性质及控制收敛定理即得 (1) 式.

5. 推论 给定概率空间 $(\Omega,\mathcal{F},\mu)$, 设 $\int f$ 存在，$A_n\in\mathcal{F}$, $n\in\mathbb{N}$, 两两不交且 $\bigcup_{n=1}^{\infty}A_n=A$, 则

$$\int_A f=\sum_{n=1}^{\infty}\int_{A_n}f. \tag{1}$$

注 由下式定义的 $\varphi: \mathcal{F} \mapsto \overline{\mathbb{R}}$ (当 f 为实函数时，而当 f 为复函数时，$\varphi: \mathcal{F} \mapsto \overline{\mathbb{C}}$)

$$\varphi(A) := \int_A f\, d\mu, \qquad A \in \mathcal{F} \tag{2}$$

是 $\mathcal{F}$ 上的一个集函数，称为 f 在 $\mathcal{F}$ 上对 μ 的 **不定积分**. 此推论断言：$(\Omega, \mathcal{F}, \mu)$ 上积分存在的函数 f 的不定积分是 $\sigma-$ 可加的集函数.

证明 设 f 为实 $\mathcal{F}-$ 可测函数，则

$$f^{\pm} I_A = \sum_{n=1}^{\infty} f^{\pm} I_{A_n}.$$

由推论 4 知

$$\begin{aligned} \int_A f^+\, d\mu &= \sum_{n=1}^{\infty} \int_{A_n} f^+\, d\mu, \\ \int_A f^-\, d\mu &= \sum_{n=1}^{\infty} \int_{A_n} f^-\, d\mu. \end{aligned} \tag{3}$$

再由 $\int f$ 存在知 (3) 中的两个级数中必有一个有限，因而两级数可以逐项相减，由此即得 (1). 对于复值函数可以应用实函数情形的结论于它的虚部和实部而得到 (1). □

6. 推论 若 f 可积，则

$$\lim_{\mu(A)\to 0} \int_A |f|\, d\mu = 0.$$

证明 令 $f_n := |f| I_{\{|f| \leqslant n\}}$, $n \in \mathbb{N}$. 则 f_n 是可测函数，且 $0 \leqslant f_n \uparrow |f|$. 于是由单调收敛定理知 $\lim_{n\to\infty} \int f_n = \int |f|$. 即对一切 $\varepsilon > 0$, 存在 $n \in \mathbb{N}$ 使

$$\int |f| I_{\{|f|>n\}} = \int |f| - \int f_n < \frac{\varepsilon}{2},$$

于是当 $\mu(A) < \varepsilon/2n$ 时

$$\int_A |f| = \int_{A\cap\{|f|\leqslant n\}} |f| + \int_{A\cap\{|f|>n\}} |f|$$
$$\leqslant n\mu(A) + \frac{\varepsilon}{2} < \varepsilon. \quad \square$$

由数学分析的方法知，下列命题成立.

7. 引理 (T,d) 是一距离空间，$a \in \overline{\mathbb{R}}$, $\{a_t:\ t \in T\} \subset \mathbb{R}$, 则对一切 $t_0 \in T$,

(1) $$\lim_{T\ni t\to t_0} a_t = a$$

的充分必要条件是对任意的 $\{t_n\} \subset T$, $t_n \to t_0(n\to\infty)$

(2) $$\lim_{n\to\infty} a_{t_n} = a.$$

由这一引理，单调收敛定理，Fatou 引理及控制收敛定理的参数 $n\to\infty$ 都可以换成在任意集 $T\subset\mathbb{R}$ 内取值的参数 $t \to t_0(t_0 \in \overline{\mathbb{R}})$, 即由离散参数换成连续参数的情形. 例如连续参数情形的控制收敛定理就是

8. 推论 设 (T,d) 是一距离空间，$(\Omega,\mathcal{F},\mu)$ 是测度空间，$\{f_t:\ t\in T\}$ 是一 $\mathcal{F}-$ 可测函数族，若存在 g 可积函数使 $|f_t| \leqslant g$, a.e., 存在 $f:\ \Omega \mapsto \overline{\mathbb{R}}$ 使 $\lim_{T\ni t\to t_0} f_t = f$, a.e. (未必 $= f_{t_0}$, a.e.), 则 $\int |f_t - f| \to 0(t\to t_0)$. 因而有

$$\lim_{t\to t_0}\int f_t = \int f.$$

证明 因为 f 牵涉到不可数个指标，因此我们需要证明 f a.e. 可测，(由于 $\left\{\sup\limits_{d(t,t_0)<\delta} f_t \leqslant x\right\} = \bigcup\limits_{d(t,t_0)<\delta}\{\omega\in\Omega:\ f_t(\omega)\leqslant x\}$ 就未必可测，或令 $\lim_{t_n\to t_0} f_{t_n} = f'$, a.e., $\lim_{s_n\to t_0} f_{s_n} = f''$, a.e., 此时 f', f'' 都 a.e. 可测，但很可能不等，只是 $f' = f''$, a.e., 而满足 $t_n \to$

t_0 的序列 $\{t_n\}\subset T$ 有不可数多个，所以 $\{\omega: \lim_{t\to t_0} f_t(\omega)$ 存在$\}$ 完全可能是一不可测集).

由假设知存在 $N\in\mathcal{F}$, $\mu(N)=0$. 当 $\omega\in N^c$ 时，$\lim_{t\to t_0} f_t(\omega) = f(\omega)$, 特别对任一给定的 $\{t_n\}\subset T$, $t_n\to t_0$. $\lim_{t_n\to t_0} f_{t_n}(\omega) = f(\omega)$, 对一切 $\omega\in N^c$. 于是 $\lim_{t_n\to t_0} f_{t_n}$ 是 a.e. 可测的且与 f a.e. 相等，故 f 是 a.e. 可测的.

$\int|f_n-f|\to 0$ 为引理 7 及定理 3 的直接推论.

利用推论 8, 可以得出一些关于积分号下微分，积分的结论.

9. 推论 设 $\delta>0$, $(t_0-\delta, t_0+\delta)\subset T\subset\mathbb{R}$, 设对一切 $t\in T$, f_t 可积，对一切 $\omega\in\Omega$, 将 $f_t(\omega)$ 看作 t 的函数，$\frac{df_t(\omega)}{dt}\Big|_{t=t_0}$ 存在，且当 $|t-t_0|<\delta$ 时，

$$\left|\frac{f_t(\omega)-f_{t_0}(\omega)}{t-t_0}\right|\leqslant g(\omega),\qquad \forall\,\omega\in\Omega, \tag{1}$$

g 可积，则

$$\left[\frac{d}{dt}\int f_t\right]_{t=t_0}=\int\left[\frac{df_t(\omega)}{dt}\right]_{t=t_0}\mu(d\omega). \tag{2}$$

证明 因为由导数定义

$$\frac{f_t(\omega)-f_{t_0}(\omega)}{t-t_0}\to\left[\frac{df_t(\omega)}{dt}\right]_{t=t_0},\quad (t\to t_0),$$

于是由推论 8 及假设 (1) 知

$$\lim_{t\to t_0}\int\frac{f_t(\omega)-f_{t_0}(\omega)}{t-t_0}\mu(d\omega)=\int\left[\frac{df_t(\omega)}{dt}\right]_{t=t_0}\mu(d\omega), \tag{3}$$

而 (3) 式左端

$$=\lim_{t\to t_0}\frac{\int f_t(\omega)\mu(d\omega)-\int f_{t_0}(\omega)\,\mu(d\omega)}{t-t_0}=\left[\frac{d}{dt}\int f_t\right]_{t=t_0}, \tag{4}$$

故由 (3), (4) 知 (2) 成立. □

10. 推论 若在区间 $[a,b]$ 上，$\int f_t$ 有限，对一切 $\omega\in\Omega$，$\frac{df_t(\omega)}{dt}$ 存在且 $\left|\frac{df_t(\omega)}{dt}\right|\leqslant g(\omega)$，$g$ 可积，则在 $[a,b]$ 上

$$(1)\qquad \frac{d}{dt}\int f_t=\int\frac{df_t(\omega)}{dt}\,\mu(d\omega).$$

证明 当 f_t，$t\in[a,b]$，都是实函数时，对一切 $\omega\in\Omega$，应用微分中值公式，存在 $t'=t'(\omega)$ 在 t_0 与 t 之间，有

$$f_t(\omega)-f_{t_0}(\omega)=(t-t_0)\left[\frac{df_t(\omega)}{dt}\right]_{t=t'}.$$

因而对一切 $t_0\in[a,b]$，$\omega\in\Omega$，

$$\left|\frac{f_t(\omega)-f_{t_0}(\omega)}{t-t_0}\right|=\left|\left[\frac{df_t(\omega)}{dt}\right]_{t=t'}\right|\leqslant g(\omega),$$

由推论 9 知 (1) 对一切 $t_0=t\in[a,b]$ 成立，亦即推论成立.

11. 推论 给定 $(\Omega,\mathcal{F},\mu)$，设 $[a,b]$ 为有限区间，$\{f_t:t\in[a,b]\}$ 为 $\mathcal{F}$ 可测函数族，对一切 $\omega\in\Omega$，$f_t(\omega)$，$t\in[a,b]$，作为 t 的函数连续，且对一切 $t\in[a,b]$，$\omega\in\Omega$，$|f_t(\omega)|\leqslant g(\omega)$，$g$ 可积，则对一切 $t\in[a,b]$

$$(1)\qquad \int_a^t\left(\int f_s\right)ds=\int\left(\int_a^t f_s(\omega)\,ds\right)\mu(d\omega).$$

又若以上的假定在每一有限区间上都成立，且 $\int_{-\infty}^{\infty}|f_t(\omega)|dt\leqslant h(\omega)$，$h$ 对 μ 可积，则

$$(2)\qquad \int_{-\infty}^{\infty}\left(\int f_t\right)dt=\int\left(\int_{-\infty}^{\infty}f_t(\omega)dt\right)\mu(d\omega).$$

证明 由控制收敛定理知 $\int f_t$ 是 t 的连续函数，所以 (1) 式左端在 $[a,b]$ 上的导数是 $\int f_t$. 而且 (1) 两边在 $t=a$ 时都为零，所以只要能证明 (1) 式右端在 $[a,b]$ 上的导数也是 $\int f_t$，即证得 (1) 式. 为此应用推论 10(将 (1) 式中的 $\int_a^t f_s(\cdot)\,ds$ 看成是推论 10 中

的 f_t), 由假设知对一切 $\omega \in \Omega$, $\frac{d}{dt}\int_a^t f_s(\omega)\,ds = f_t(\omega)$ 存在, 且 $\left|\frac{d}{dt}\int_a^t f_s(\omega)\,ds\right| \leqslant g(\omega)$ g 可积, 又

$$\left|\int\left(\int_a^t f_s(\omega)\,ds\right)\mu(d\omega)\right| \leqslant \int\left|\int_a^t f_s(\omega)\,ds\right|\mu(d\omega)$$
$$\leqslant \int (t-a)|g(\omega)|\,\mu(d\omega) < \infty,$$

故由推论 10 知

$$\frac{d}{dt}\int\left(\int_a^t f_s(\omega)\,ds\right)\mu(d\omega) = \int\left(\frac{d}{dt}\int_a^t f_s(\omega)\,ds\right)\mu(d\omega)$$
$$= \int f_t(\omega)\,\mu(d\omega),$$

故 (1) 式获证.

12. 我们这里所讲的抽象空间上的积分不仅包括 L- 积分, r.v. 的数学期望作为特殊例子, 而且包括了无穷级数作为特例, 事实上, 令 $\Omega = \mathbb{N}$, $\mathcal{F}$ 为 Ω 的一切子集类, 定义 $\mu(\{n\}) = 1$, $n \in \mathbb{N}$, $\mu(A) = |A|$(即 A 的元数). 则 $(\mathbb{N}, \mathcal{F}, \mu)$ 为一 σ- 有限测度空间 (通常称此测度为 $\mathbb{N}$ 上的计数测度). 则在此空间上定义的函数 f 为一个序列 $\{a_n, n \in \mathbb{N}\}$, 即 $f(n) = a_n$, $n \in \mathbb{N}$, 所以 $\int f\,d\mu = \sum_{n=1}^{\infty} a_n$ 为一无穷级数, 因此将积分收敛定理应用于这种情形很容易得到一些关于无穷级数的逐项取极限的结论, 例如:

1) 当 $m \to \infty$ 时, $0 \leqslant a_{nm} \uparrow a_n$, 对一切 $n \in \mathbb{N}$ 或 $|a_{nm}| \leqslant b_n$, 对一切 $n, m \in \mathbb{N}$, $\sum_{n=1}^{\infty} b_n < \infty$ 且 $\lim_{m\to\infty} a_{nm} = a_n$, 对一切 $n \in \mathbb{N}$, 则

$$\lim_{m\to\infty}\sum_{n=1}^{\infty} a_{nm} = \sum_{n=1}^{\infty} a_n.$$

2) 若 $a_{nm} \geqslant 0$, 对一切 $n, m \in \mathbb{N}$, 则

$$\varliminf_{m\to\infty}\sum_{n=1}^{\infty} a_{nm} \geqslant \sum_{n=1}^{\infty}\varliminf_{m\to\infty} a_{nm}.$$

自然此处的参数 m 还可以一般化，也可以得到一些逐项微分，逐项积分的结论.

作为推论 5 及 12.1) 的一个应用，我们证明下面的：

13. 引理 设 X 为任一 r.v.，则

$$\sum_{n=1}^{\infty}\mathbf{P}(|X|\geqslant n)\leqslant \mathbf{E}|X|\leqslant 1+\sum_{n=1}^{\infty}\mathbf{P}(|X|\geqslant n).$$

证明 令 $\Lambda_n:=\{n\leqslant |X|<n+1\}$，则

$$\sum_{n=0}^{\infty}n\mathbf{P}(\Lambda_n)\leqslant \mathbf{E}|X|=\sum_{n=0}^{\infty}\int_{\Lambda_n}|X|\,d\mathbf{P}$$
$$\leqslant\sum_{n=0}^{\infty}(n+1)\mathbf{P}(\Lambda_n)=1+\sum_{n=0}^{\infty}n\mathbf{P}(\Lambda_n),$$

于是只要证

$$\sum_{n=0}^{\infty}n\mathbf{P}(\Lambda_n)=\sum_{n=1}^{\infty}\mathbf{P}(|X|\geqslant n).$$

事实上，

$$\sum_{n=0}^{N}n\mathbf{P}(\Lambda_n)=\sum_{n=1}^{N}\sum_{k=1}^{n}\mathbf{P}(n\leqslant |X|<n+1)$$
$$=\sum_{k=1}^{N}\sum_{n=k}^{N}\mathbf{P}(n\leqslant |X|<n+1)=\sum_{k=1}^{N}\mathbf{P}(k\leqslant |X|<N+1).$$

令

$$a_{kN}=\begin{cases}\mathbf{P}(k\leqslant |X|<N+1), & 1\leqslant k\leqslant N,\\ 0, & k>N,\end{cases}$$

则当 $N\uparrow\infty$ 时， $a_{kN}\uparrow\mathbf{P}(|X|\geqslant k)$, 由 12.1) 的单调收敛部分即得

$$\sum_{n=0}^{\infty}n\mathbf{P}(\Lambda_n)=\lim_{N\to\infty}\sum_{n=0}^{N}n\mathbf{P}(\Lambda_n)$$

$$=\lim_{N\to\infty}\sum_{k=1}^{\infty}a_{kN}=\sum_{k=1}^{\infty}\mathbf{P}(|X|\geqslant k).\quad\square$$

习题

1. (Fatou 引理的推广) 设 U_n, V_n 可积, 且 $U_n\xrightarrow{\text{a.e.}}U$、$V_n\xrightarrow{\text{a.e.}}V$, $\int U_n\to\int U$ 有限， $\int V_n\to\int V$ 有限.

1) 若 $U_n\leqslant X_n$, 对一切 $n\in\mathbb{N}$, 则

$$\int\varliminf_{n\to\infty}X_n\leqslant\varliminf_{n\to\infty}\int X_n$$

2) 若 $X_n\leqslant V_n$, 对一切 $n\in\mathbb{N}$, 则

$$\varlimsup_{n\to\infty}\int X_n\leqslant\int\varlimsup_{n\to\infty}X_n$$

2. (控制收敛定理的推广) 设 $|X_n|\leqslant U_n$ 而 U_n 可积 $U_n\xrightarrow{\text{a.e.}}U$ 且 $\int U_n\to\int U$ 有限，则当 $X_n\xrightarrow{\text{a.e.}}X$ 时，有

$$\int|X_n-X|\to 0,\qquad \text{因而}\qquad \int X_n\to\int X.$$

3. 试证：给定具有有限期望的随机变量 X 及 $\varepsilon>0$, 存在一个简单函数 X_ε, 使得

$$\mathbf{E}|X-X_\varepsilon|<\varepsilon,\qquad |X_\varepsilon|\leqslant|X|.$$

因而存在一个简单函数序列 $\{X_m\}$, 使得对一切 $m\in\mathbb{N}$, $|X_m|\leqslant|X|$, 且

$$\lim_{m\to\infty}\mathbf{E}|X-X_m|=0.$$

4. 对于 $\mathscr{F}$ 中任何两个集 Λ_1 与 Λ_2, 定义

$$\rho(\Lambda_1,\Lambda_2)=\mathbf{P}(\Lambda_1\Delta\Lambda_2),$$

则 ρ 是 $\mathcal{F}$ 中的集的空间中的伪度量 (即除 $\rho(\Lambda_1, \Lambda_2) = 0 \Rightarrow \Lambda_1 = \Lambda_2$ 外，距离的其它假设都满足); 称引入 ρ 后的空间 $\mathcal{F}$ 为度量空间 $M(\mathcal{F}, \rho)$, 证明对于每个可积的随机变量 X, 由 $\Lambda \to \int_\Lambda X\,d\mathbf{P}$ 给出的由 $M(\mathcal{F}, \rho)$ 到 $\mathbb{R}^1$ 的映射是连续的. 类似地，由

$$(\Lambda_1, \Lambda_2) \to \Lambda_1 \cup \Lambda_2, \quad \Lambda_1 \cap \Lambda_2, \quad \Lambda_1 \setminus \Lambda_2, \quad \Lambda_1 \Delta \Lambda_2.$$

给出的由 $M(\mathcal{F}, \rho) \times M(\mathcal{F}, \rho)$ 到 $M(\mathcal{F}, \rho)$ 的映射都是连续的. 如果去掉一个零概率集后,

$$\limsup_n \Lambda_n = \liminf_n \Lambda_n,$$

则我们用 $\lim_n \Lambda_n$ 表示这两个集共同的等价类. 证明在这种情况下 $\{\Lambda_n\}$ 按度量 ρ 收敛于 $\lim_n \Lambda_n$. 作为一个特殊情况，试推出：

如果 $\mathbf{E}|X| < \infty$, 且 $\lim_n \mathbf{P}(\Lambda_n) = 0$, 则 $\lim_{n\to\infty} \int_{\Lambda_n} X\,d\mathbf{P} = 0$, 特别有

$$\lim_{n\to\infty} \int_{\{|X|>n\}} X\,d\mathbf{P} = 0.$$

第六章 乘积测度与无穷乘积概率空间

在数学分析中，低维空间与高维空间的表示法、体积以及积分之间的关系，例如重积分化累次积分都是很重要的．本章我们将讨论抽象空间中的这些问题，即如何由"低维"可测空间及其测度构造"高维"可测空间及其测度；如何由"重积分"化"累次积分"以及如何构造任意多个概率空间的乘积概率空间的问题．

§1. 乘积测度与转移测度

1. 定义 给定集 A_1, A_2, 称集

(1)
$$A_1 \times A_2 := \{(\omega_1, \omega_2):\ \omega_i \in A_i,\ i = 1, 2\}$$

为 A_1, A_2 的**乘积集**，容易看出乘积集有以下性质：

(2)
$$A_1 \times A_2 = \varnothing \iff A_1 = \varnothing \text{ 或} A_2 = \varnothing$$

若 $A_1 \times A_2 \neq \varnothing$, 则

(3)
$$A_1 \times A_2 \subset B_1 \times B_2 \iff A_i \subset B_i,\ i = 1, 2,$$
$$A_1 \times A_2 = B_1 \times B_2 \iff A_i = B_i,\ i = 1, 2.$$

若 I 是任意指标集，则

(4)
$$\bigcap_{\alpha \in I} (A_{1\alpha} \times A_{2\alpha}) = \Big(\bigcap_{\alpha \in I} A_{1\alpha}\Big) \times \Big(\bigcap_{\alpha \in I} A_{2\alpha}\Big),$$

特别

$$\bigcap_{\alpha \in I} (A_1 \times A_{2\alpha}) = A_1 \times \Big(\bigcap_{\alpha \in I} A_{2\alpha}\Big),$$
$$\bigcap_{\alpha \in I} (A_{1\alpha} \times A_2) = \Big(\bigcap_{\alpha \in I} A_{1\alpha}\Big) \times A_2.$$

若 I 是任意指标集,

$$(5)\qquad \begin{cases} \bigcup\limits_{\alpha\in I}(A_{1\alpha}\times A_2)=\Big(\bigcup\limits_{\alpha\in I}A_{1\alpha}\Big)\times A_2,\\ \bigcup\limits_{\alpha\in I}(A_1\times A_{2\alpha})=A_1\times\Big(\bigcup\limits_{\alpha\in I}A_{2\alpha}\Big).\end{cases}$$

若 $A_1\times A_2\subset B_1\times B_2$, 则

$$(6)\qquad \begin{aligned}(B_1\times B_2)\setminus(A_1\times A_2)=&((B_1\setminus A_1)\times A_2)\\ &\cup(A_1\times(B_2\setminus A_2))\cup((B_1\setminus A_1)\times(B_2\setminus A_2)).\end{aligned}$$

而且上式右边各集两两不交.

2. 引理 设 Ω_1, Ω_2 是给定的两个空间 (非空集合), $\mathcal{S}_i$ 是 Ω_i 的半集代数, $i=1,2$, 则

(i) $\mathcal{S}:=\{A_1\times A_2:\ A_i\in\mathcal{S}_i,\ i=1,2\}$ 是 $\Omega_1\times\Omega_2$ 中的半集代数.

(ii) $\mathcal{A}:=\left\{\bigcup_{i=1}^n A_i:\ n\geq 1,\ A_1,\cdots,A_n\in\mathcal{S}, \text{两两不交}\right\}$ 是 $\Omega_1\times\Omega_2$ 中包含 $\mathcal{S}$ 的最小集代数.

证明 由 (1.4), (1.6) 即知 $\Omega_1\times\Omega_2\in\mathcal{S}$, $\mathcal{S}$ 对交封闭, 且它的任一集的余集可表为 $\mathcal{S}$ 的有限个两两不交的集的并, 故 $\mathcal{S}$ 为 $\Omega_1\times\Omega_2$ 的一个半集代数.

由引理 3.1.6 及 (i) 知 $\mathcal{A}$ 为 $\Omega_1\times\Omega_2$ 中包含 $\mathcal{S}$ 的最小集代数.
□

3. 定义 设 $(\Omega_i,\mathcal{F}_i)$, $i=1,2$ 为两个可测空间, 称包含集类

$$\mathcal{C}:=\{A_1\times A_2\ A_i\in\mathcal{F}_i,\ i=1,2\}$$

的最小 $\sigma-$ 代数 $\sigma(\mathcal{C})$ 为 $\mathcal{F}_1,\mathcal{F}_2$ 的 **乘积 $\sigma-$ 代数**, 记作 $\mathcal{F}_1\times\mathcal{F}_2$.

为叙述方便, 称 $A_1\times A_2$, $A_i\subset\Omega_i$, $i=1,2$, 为 $\Omega_1\times\Omega_2$ 中的**矩形**, 而 $\mathcal{C}$ 中的元称为 $\Omega_1\times\Omega_2$ 中的 **可测矩形**.

例1　设 $\Omega_1=\Omega_2=\mathbb{R}$, $\mathcal{F}_1=\mathcal{F}_2=\mathcal{B}$, 则 $\mathbb{R}^2=\mathbb{R}\times\mathbb{R}$, $\mathcal{B}^2=\mathcal{B}\times\mathcal{B}$.

事实上，对一切 $x=(x_1,x_2)\in\mathbb{R}^2$,

$$(-\infty,x]=(-\infty,x_1]\times(-\infty,x_2]\in\mathcal{B}\times\mathcal{B}.$$

而 $\mathcal{B}^2=\sigma(\{(-\infty,x]:\ x\in\mathbb{R}^2\})$, 故 $\mathcal{B}^2\subset\mathcal{B}\times\mathcal{B}$. 反之，只需证：

$$A_1\times A_2\in\mathcal{B}^2,\quad\forall\ A_i\in\mathcal{B},\quad i=1,2. \tag{1}$$

对任意给定的 $a_2\in\mathbb{R}$, 令

$$\mathfrak{M}_1:=\{A_1\in\mathcal{B}:\ A_1\times(-\infty,a_2]\in\mathcal{B}^2\},$$

则易证：$\mathfrak{M}_1$ 是 $\mathbb{R}$ 中的 $\sigma-$ 代数，且包含 $\{(-\infty,a_1]:\ a_1\in\mathbb{R}\}$, 所以 $\mathfrak{M}_1=\mathcal{B}$, 即：

$$A_1\times(-\infty,a_2]\in\mathcal{B}^2,\quad\forall\ A_1\in\mathcal{B},\quad a_2\in\mathbb{R}. \tag{2}$$

其次，对任给的 $A_1\in\mathcal{B}$, 考查

$$\mathfrak{M}_2:=\{A_2\in\mathcal{B}:\ A_1\times A_2\in\mathcal{B}^2\}.$$

则易证：$\mathfrak{M}_2$ 是包含 $\{(-\infty,a_2]:\ a_2\in\mathbb{R}\}$ 的 $\sigma-$ 代数，所以 $\mathfrak{M}_2=\mathcal{B}$. 这就得到 (1), 因而

$$\mathcal{B}\times\mathcal{B}=\sigma\left(\{A_1\times A_2:\ A_i\in\mathcal{B},\ i=1,2\}\right)\subset\mathcal{B}^2.$$

故 $\mathcal{B}^2=\mathcal{B}\times\mathcal{B}$.

例2　$\Omega_1=\Omega_2=\overline{\mathbb{R}}$, $\mathcal{F}_1=\mathcal{F}_2=\overline{\mathcal{B}}$, 则 $\overline{\mathbb{R}}\times\overline{\mathbb{R}}=\overline{\mathbb{R}}^2$, $\overline{\mathcal{B}}\times\overline{\mathcal{B}}=\overline{\mathcal{B}}^2$.

例3　$\Omega_1=\mathbb{R}^n$, $\Omega_2=\mathbb{R}^m$, $\mathcal{F}_1=\mathcal{B}^n$, $\mathcal{F}_2=\mathcal{B}^m$, 则 $\mathbb{R}^{n+m}=\mathbb{R}^n\times\mathbb{R}^m$, $\mathcal{B}^{n+m}=\mathcal{B}^n\times\mathcal{B}^m$, $n,m\in\mathbb{N}$.

例4　$\Omega_1=\overline{\mathbb{R}}^n$, $\Omega_2=\overline{\mathbb{R}}^m$, $\mathcal{F}_1=\overline{\mathcal{B}}^n$, $\mathcal{F}_2=\overline{\mathcal{B}}^m$, 则 $\overline{\mathbb{R}}^{n+m}=\overline{\mathbb{R}}^n\times\overline{\mathbb{R}}^m$, $\overline{\mathcal{B}}^{n+m}=\overline{\mathcal{B}}^n\times\overline{\mathcal{B}}^m$, $n,m\in\mathbb{N}$.

例5　设 $\Omega_1=\mathfrak{C}^n$, $\Omega_2=\mathfrak{C}^m$, 其中 $\mathfrak{C}:=\{z:\ z=x+iy,\ x,y\in\mathbb{R}\}$ 表示全体复数集，$\mathcal{F}_1=\mathcal{B}_z^n$, $\mathcal{F}_2=\mathcal{B}_z^m$, 则 $\mathfrak{C}^{n+m}=\mathfrak{C}^n\times\mathfrak{C}^m$, $\mathcal{B}_z^{n+m}=\mathcal{B}_z^n\times\mathcal{B}_z^m$, $n,m\in\mathbb{N}$.

例 2— 例 4 的证法与例 1 同，请读者作为练习证明例 5.

稍微留意一下. 如果 $\mathbb{R}^n$ 中的点采用 $(x_1,\cdots,x_n)$ 的记法, 那么由 (1.1) 式 $\mathbb{R}^{n+m}$ 和 $\mathbb{R}^n\times\mathbb{R}^m$ 在形式上是不一样的, 它们的元分别是 $(x_1,\cdots,x_{n+m})$, $((x_1,\cdots,x_n),(x_{n+1},\cdots,x_{n+m}))$, 为此我们给出如下的

4. 定义 设 A_k. $k=1,2,\cdots,n$ 是 n 个集合, 则集合

$$\prod_{k=1}^n A_k := A_1\times\cdots\times A_n :$$
$$=\big\{(\omega_1,\cdots,\omega_n):\ \omega_k\in A_k,\, k=1,2,\cdots,n\big\}.$$

称为 $A_1,\cdots,A_n$ 的 **乘积集**.

设 $(\Omega_k,\mathcal{F}_k)$, $k=1,2,\cdots,n$, 是 n 个可测空间, $A_k\subset\Omega_k(\in\mathcal{F}_k)$, $k=1,2,\cdots,n$, 则 $A_1\times\cdots\times A_n$ 称为 $\Omega_1\times\cdots\times\Omega_n$ 的以 $A_1,\cdots,A_n$ 为边的 **矩形** (**可测矩形**). $\Omega_1\times\cdots\times\Omega_n$ 中一切可测矩形类生成的 $\sigma-$ 代数称为 $\mathcal{F}_k$, $k=1,2,\cdots,n$, 的 **乘积 $\sigma-$ 代数**, 记作 $\prod_{k=1}^n\mathcal{F}_k:=\mathcal{F}_1\times\cdots\times\mathcal{F}_n$, 即

$$\mathcal{F}_1\times\cdots\times\mathcal{F}_n=\sigma\left(\{A_1\times\cdots A_n:\ A_k\in\mathcal{F}_k,\ k=1,2,\cdots,n\}\right).$$

关于两个集的乘积集的性质 (1.2) – (1.6) 以及引理 2 都可以推广至 n 个集的乘积集的情形, 这一点留作习题.

乘积集运算是不满足结合律的, 例如 $(A_1\times A_2)\times A_3$, $A_1\times(A_2\times A_3)$, $A_1\times A_2\times A_3$ 的元素为 $((\omega_1,\omega_2),\omega_3)$, $(\omega_1,(\omega_2,\omega_3))$, $(\omega_1,\omega_2,\omega_3)$ 是不相同的, 但是它们的元素构成是一样的 (都是 ω_1, ω_2,ω_3, 而且顺序也是一样的). 只是构作的过程不同, 而对于我们所讨论的问题来说, 这是不重要的. 因此我们今后对它们不加区别, 即对 $A_1,\cdots,A_n$ 来说, 我们认为 $A_1,\cdots,A_n$ 的乘积集运算任意结合 (只要顺序不变) 而得到的乘积都是一样的, n 个集的乘积集也是这样, 于是 $\mathbb{R}^n=\underbrace{\mathbb{R}\times\cdots\times\mathbb{R}}_{n}$, 而 $\mathbb{R}^{n+m}=\mathbb{R}^n\times\mathbb{R}^m$ 就是合理的.

对于 n 个 $\sigma-$ 代数的乘积 $\sigma-$ 代数运算我们也作这样的约定，于是 $\mathcal{B}^n := \mathcal{B}\times\cdots\times\mathcal{B}$, 而 $\mathcal{B}^{n+m} = \mathcal{B}^n\times\mathcal{B}^m$ 也就合理了.

为了更易理解，我们对乘积 $\sigma-$ 代数上的测度，可测函数及积分问题的讨论集中于两个 $\sigma-$ 代数的乘积 $\sigma-$ 代数的情形. 事实上, 推广到有限维的情形没有本质的困难, 对于一般有限维的情形的结果，我们将述而不证.

5. 定义 设 $A\subset\Omega_1\times\Omega_2$. 对一切 $\omega_1\in\Omega_1$、$\omega_2\in\Omega_2$, 称集合

(1) $$A_{\omega_1} := A(\omega_1) := \left\{\tilde{\omega}_2\in\Omega_2:\ (\omega_1,\tilde{\omega}_2)\in A\right\},$$

(2) $$A_{\omega_2} := A(\omega_2) := \left\{\tilde{\omega}_1\in\Omega_1:\ (\tilde{\omega}_1,\omega_2)\in A\right\}.$$

分别为 A **在** ω_1,ω_2 **处的截集**. 显然 $A_{\omega_1}\subset\Omega_2$, $A_{\omega_2}\subset\Omega_1$, 且 $A=\emptyset \implies \forall\,\omega_i\in\Omega_i$, $A(\omega_i)=\emptyset$, $i=1,2$. 此外，容易验证具有下列诸性质：对一切 $\omega_i\in\Omega_i$, $i=1,2$, 有

(3) $$A\cap B=\emptyset \Longrightarrow A(\omega_i)\cap B(\omega_i)=\emptyset;$$

(4) $$A\subset B \Longrightarrow A(\omega_i)\subset B(\omega_i);$$

(5) $$A=\bigcup_\alpha A^{(\alpha)} \Longrightarrow A(\omega_i)=\bigcup_\alpha A^{(\alpha)}(\omega_i);$$

(6) $$A=\bigcap_\alpha A^{(\alpha)} \Longrightarrow A(\omega_i)=\bigcap_\alpha A^{(\alpha)}(\omega_i);$$

(7) $$C=A\setminus B \Longrightarrow C(\omega_i)=A(\omega_i)\setminus B(\omega_i);$$

特别

(8) $$(A^c)(\omega_i)=(A(\omega_i))^c.$$

6. 定理 设 $(\Omega_i,\mathcal{F}_i)$, $i=1,2$, 是可测空间. 若 $A\in\mathcal{F}_1\times\mathcal{F}_2$, 则对一切 $\omega_1\in\Omega_1$, $A(\omega_1)\in\mathcal{F}_2$, 而对一切 $\omega_2\in\Omega_2$, $A(\omega_2)\in\mathcal{F}_1$.

这条定理可以简单地叙述 (虽然不算确切, 但很形象) 为：$\mathcal{F}_1\times\mathcal{F}_2$ 可测集的任何截集是可测集.

证明 令

$$\mathfrak{M} := \left\{A \in \mathcal{F}_1 \times \mathcal{F}_2 : \ \forall\, \omega_i \in \Omega_i, \quad A(\omega_i) \in \mathcal{F}_j, \quad i \neq j, i, j = 1, 2\right\},$$

$$\mathcal{C} := \left\{A_1 \times A_2 : \ A_i \in \mathcal{F}_i, \quad i = 1, 2\right\},$$

则由

$$(A_1 \times A_2)_{\omega_1} = \begin{cases} A_2, & \omega_1 \in A_1, \\ \varnothing, & \omega_1 \notin A_1, \end{cases}$$

$$(A_1 \times A_2)_{\omega_2} = \begin{cases} A_1, & \omega_2 \in A_2, \\ \varnothing, & \omega_2 \notin A_2, \end{cases}$$

知 $\mathfrak{M} \supset \mathcal{C}$. 因而 $\Omega_1 \times \Omega_2 \in \mathfrak{M}$. 再由 (5.5) 和 (5.8) 知 $\mathfrak{M}$ 对可数并及求余封闭. 所以 $\mathfrak{M}$ 是一 $\sigma-$ 代数. 故

$$\mathcal{F}_1 \times \mathcal{F}_2 \supset \mathfrak{M} \supset \sigma(\mathcal{C}) = \mathcal{F}_1 \times \mathcal{F}_2.$$

即 $\mathfrak{M} = \mathcal{F}_1 \times \mathcal{F}_2$, 定理获证. □

7. 定义 设 $f = f(\cdot, \cdot)$ 是定义在 $\Omega_1 \times \Omega_2$ 上的函数 (实值或复值). 对任意给定的 $\omega_1 \in \Omega_1$, Ω_2 上的函数

$$f_{\omega_1} := f(\omega_1, \cdot)$$

称为 **f 在 ω_1 的截函数**. 同样, 对任意给定的 $\omega_2 \in \Omega_2$, Ω_1 上的函数

$$f_{\omega_2} := f(\cdot, \omega_2)$$

称为 **f 在 ω_2 的截函数**.

8. 定理 给定 $(\Omega_i, \mathcal{F}_i)$, $i = 1, 2$. 若 f 是实值 (或复值) $\mathcal{F}_1 \times \mathcal{F}_2$ 可测函数, 则对一切 $\omega_i \in \Omega_i$, f_{ω_i} 是 $\mathcal{F}_j$ 可测函数, $j \neq i$, $i, j = 1, 2$. 简言之, $\mathcal{F}_1 \times \mathcal{F}_2$ 可测函数的截函数是可测的.

证明 对一切 $B \in \mathcal{B}$ (只讨论实函数的情形). 由假设知

$$f^{-1}(B) = \left\{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 : \ f(\omega_1, \omega_2) \in B\right\} \in \mathcal{F}_1 \times \mathcal{F}_2.$$

于是对给定的 $\omega_1 \in \Omega_1$,

$$\begin{aligned} f_{\omega_1}^{-1}(B) &= \{\omega_2 \in \Omega_2 : f_{\omega_1}(\omega_2) \in B\} \\ &= \{\omega_2 \in \Omega_2 : f(\omega_1, \omega_2) \in B\} \\ &= \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in f^{-1}(B)\} = (f^{-1}(B))_{\omega_1} \end{aligned}$$

由定理 6 知 $f_{\omega_1}^{-1}(B) \in \mathcal{F}_2$, 即 f_{ω_1} 为 $\mathcal{F}_2-$ 可测, 同法可证: 对给定的 $\omega_2 \in \Omega_2$, f_{ω_2} 是 $\mathcal{F}_1$ 可测的. □

9. 定理 给定 $(\Omega_i, \mathcal{F}_i)$, $i = 1, 2$, 若 f 是非负 $\mathcal{F}_1 \times \mathcal{F}_2$ 可测函数, μ_i 是 $\mathcal{F}_i$, $i = 1, 2$, 上的 $\sigma-$ 有限测度, 则

$$\begin{gathered} f^{(2)} := \int_{\Omega_1} f(\omega_1, \cdot)\mu_1(d\omega_1) \\ (\text{相应地} f^{(1)} := \int_{\Omega_1} f(\cdot, \omega_2)\mu_1(d\omega_2)) \end{gathered} \tag{1}$$

是非负 $\mathcal{F}_2$ (相应地, $\mathcal{F}_1-$) 可测函数.

特别, 对一切 $A \in \mathcal{F}_1 \times \mathcal{F}_2$, $\mu_1(A(\omega_2))$ (相应地, $\mu_2(A(\omega_1))$) 作为 $\Omega_2(\Omega_1)$ 上的函数是非负 $\mathcal{F}_2(\mathcal{F}_1)-$ 可测的.

证明 由于证法相同, 只证 $f^{(2)}$ 是 $\mathcal{F}_2-$ 可测的. 若此事获证, 则取 $f = I_A$, $A \in \mathcal{F}_1 \times \mathcal{F}_2$. 由 $A(\omega_2)$ 的定义知

$$\begin{aligned} \mu_1(A(\omega_2)) &= \int_{\Omega_1} I_{A(\omega_2)}(\omega_1)\,\mu_1(d\omega_1) \\ &= \int_{\Omega_1} I_A(\omega_1, \omega_2)\,\mu_1(d\omega_1). \end{aligned} \tag{2}$$

因而后一结论获证.

1) 先用 $\lambda-$ 系方法证明: 对任何 $A \in \mathcal{F}_1 \times \mathcal{F}_2$,

$$\int_{\Omega_1} I_A(\omega_1, \cdot)\,\mu_1(d\omega_1)$$

可测.

注意 μ_1 是 $\mathcal{F}_1$ 上的 $\sigma-$ 有限测度，所以存在 $\Omega_1^{(n)} \in \mathcal{F}_1, n \in \mathbb{N}$ 两两不交， $\mu_1(\Omega_1^{(n)}) < \infty$、$n \in \mathbb{N}$，且 $\bigcup_{n\in\mathbb{N}} \Omega_1^{(n)} = \Omega_1$. 令

$$\mathfrak{M} := \Big\{A \in \mathcal{F}_1 \times \mathcal{F}_2 : \ \forall\, n \in \mathbb{N},$$

$$\int_{\Omega_1^{(n)}} I_A(\omega_1, \cdot)\, \mu_1(d\omega_1) \text{为} \mathcal{F}_2 - \text{可测}\Big\},$$

于是，若 $A \in \mathfrak{M}$, 则

$$\int_{\Omega_1} I_A(\omega_1, \cdot)\, \mu_1(d\omega_1) = \sum_{n=1}^{\infty} \int_{\Omega_1^{(n)}} I_A(\omega_1, \cdot)\, \mu_1(d\omega_1)$$

$\mathcal{F}_2-$ 可测. 又易见

$$\mathcal{C} := \{A_1 \times A_2 : \ A_i \in \mathcal{F}_i, \ i = 1, 2\}$$

是 $\pi-$ 系，且若 $A \in \mathcal{C}$, 则 $A = A_1 \times A_2$、$A_i \in \mathcal{F}_i$, 于是，对一切 $n \in \mathbb{N}$,

$$\int_{\Omega_1^{(n)}} I_A(\omega_1, \cdot)\, \mu_1(d\omega_1) = I_{A_2} \cdot \mu(A_1 \cap \Omega_1^{(n)})$$

是 $\mathcal{F}_2-$ 可测的，所以 $\mathcal{C} \subset \mathfrak{M}$, 因此，由 $\lambda-$ 系方法，只要证 $\mathfrak{M}$ 是 $\lambda-$ 系 (因为此时就有 $\mathfrak{M} \supset \sigma(\mathcal{C}) = \mathcal{F}_1 \times \mathcal{F}_2$).

因为 $\Omega_1 \times \Omega_2 \in \mathcal{C}$, 所以 $\Omega_1 \times \Omega_2 \in \mathfrak{M}$. 其次，若 $A, B \in \mathfrak{M}$, 且 $A \subset B$, 则对一切 $n \in \mathbb{N}$, $\int_{\Omega_1^{(n)}} I_A(\omega_1, \cdot)\, \mu_1(d\omega_1)$, $\int_{\Omega_1^{(n)}} I_B(\omega_1, \cdot)\, \mu_1(d\omega_1)$ 都是 $\mathcal{F}_2-$ 可测的，且可证它们都是有限函数. 于是对一切 $n \in \mathbb{N}$,

$$\int_{\Omega_1^{(n)}} I_{B\setminus A}(\omega_1, \cdot)\, \mu_1(d\omega_1)$$

$$= \int_{\Omega_1^{(n)}} I_B(\omega_1, \cdot)\, \mu_1(d\omega_1) - \int_{\Omega_1^{(n)}} I_A(\omega_1, \cdot)\, \mu_1(d\omega_1)$$

是 $\mathcal{F}_2-$ 可测函数，因而 $B \setminus A \in \mathfrak{M}$. 再设 $A_m \in \mathfrak{M}$, $m \in \mathbb{N}$; $A_m \uparrow A$, 当 $m \uparrow$ 时，则 $\forall\, n \in \mathbb{N}$.

$$\int_{\Omega_1^{(n)}} I_A(\omega_1, \cdot)\, \mu_1(d\omega_1) = \lim_{m\to\infty} \int_{\Omega_1^{(n)}} I_{A_m}(\omega_1, \cdot)\, \mu_1(d\omega_1)$$

$\mathcal{F}_2$ 可测，因而 $A \in \mathfrak{M}$. 故 $\mathfrak{M}$ 是一 $\lambda-$ 系.

2) 由 1). 当 $f = I_A$, $A \in \mathcal{F}_1 \times \mathcal{F}_2$ 时，$f^{(2)}$ 是 $\mathcal{F}_2-$ 可测的，于是由积分的线性性质，当 f 为非负 $\mathcal{F}_1 \times \mathcal{F}_2$ 简单函数时，$f^{(2)}$ 是 $\mathcal{F}_2-$ 可测，由单调收敛定理知，当 f 为非负 $\mathcal{F}_1 \times \mathcal{F}_2$ 可测时，$f^{(2)}$ 是 $\mathcal{F}_2-$ 可测的. 故定理获证. □

10. 定理 设 $(\Omega_i, \mathcal{F}_i, \mu_i)$, $i = 1, 2$, 是两个 $\sigma-$ 有限测度空间，若令

(1) $$\mu(A) := \int_{\Omega_1} \mu_2(A(\omega_1))\, \mu_1(d\omega_1), \qquad A \in \mathcal{F}_1 \times \mathcal{F}_2,$$

或

(2) $$\mu(A) := \int_{\Omega_2} \mu_1(A(\omega_2))\, \mu_2(d\omega_2), \qquad A \in \mathcal{F}_1 \times \mathcal{F}_2,$$

则 μ 是 $\mathcal{F}_1 \times \mathcal{F}_2$ 上唯一满足：

(3) $$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2), \qquad \forall\, A_i \in \mathcal{F}_i, \quad i = 1, 2.$$

的 $\sigma-$ 有限测度.

由 (1) (或 (2)) 决定的测度称为 μ_1、μ_2 的**乘积测度**. 通常记为 $\mu_1 \times \mu_2$. $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mu_1 \times \mu_2)$ 称为 $(\Omega_i, \mathcal{F}_i, \mu_i)$, $i = 1, 2, \cdots, n$, 的**乘积测度空间**.

证明 显然由 (1) 或 (2) 定义的 μ 满足 (3).

下面先证满足 (3) 的测度 μ 的 $\sigma-$ 有限性及唯一性. 事实上，由于 μ_i 是 $\sigma-$ 有限的，所以存在 $\Omega_i^{(n)}$, $n \in \mathbb{N}$, 两两不交，$\mu_i(\Omega_i^{(n)}) < \infty$. 且 $\bigcup_{n=1}^{\infty} \Omega_i^{(n)} = \Omega_i$, $i = 1, 2$. 于是 $\Omega_1^{(m)} \times \Omega_2^{(n)}$, $m, n \in \mathbb{N}$, 两两不交，$\bigcup_{m,n=1}^{\infty} (\Omega_1^{(m)} \times \Omega_2^{(n)}) = \Omega_1 \times \Omega_2$, 且

$$\mu(\Omega_1^{(m)} \times \Omega_2^{(n)}) = \mu_1(\Omega_1^{(m)})\mu_2(\Omega_2^{(n)}) < \infty.$$

即 μ 为 $\sigma-$ 有限. 若还有一 μ' 满足 (3), 则 μ' 也是 $\sigma-$ 有限的，且对一切 $A_i \in \mathcal{F}_i$, $i = 1, 2$,

$$\mu'(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2) = \mu(A_1 \times A_2).$$

即 μ, μ' 在可测矩形集类 $\mathcal{C}$ 上相等，而 $\mathcal{C}$ 为半集代数， $\sigma(\mathcal{C}) = \mathcal{F}_1 \times \mathcal{F}_2$，故由测度扩张定理知 μ, μ' 在 $\mathcal{F}_1 \times \mathcal{F}_2$ 上一致.

再证由 (1) 定义的 μ 是 $\mathcal{F}_1 \times \mathcal{F}_2$ 上的测度. 由定理 9 知 $\mu_2(A(\omega_1))$ 作为 Ω_1 上的函数是非负 $\mathcal{F}_1-$ 可测的，因而对一切 $A \in \mathcal{F}_1 \times \mathcal{F}_2$，$\mu(A)$ 有意义且非负. 因此只需证： μ 具 $\sigma-$ 可加性.

事实上，设 $A^{(n)} \in \mathcal{F}_1 \times \mathcal{F}_2$，$n \in \mathbb{N}$，两两不交，则对一切 $\omega_1 \in \Omega_1$，$A^{(n)}(\omega_1) \in \mathcal{F}_2$，$n \in \mathbb{N}$，两两不交. 因而，由 μ_1 的 $\sigma-$ 可加性及单调收敛定理知

$$\mu\Big(\bigcup_{n=1}^{\infty} A^{(n)}\Big) = \int_{\Omega_1} \mu_2\Big(\bigcup_{n=1}^{\infty} A^{(n)}(\omega_1)\Big)\,\mu_1(d\omega_1)$$

$$= \int_{\Omega_1} \sum_{n=1}^{\infty} \mu_2\big(A^{(n)}(\omega_1)\big)\,\mu_1(d\omega_1)$$

$$= \sum_{n=1}^{\infty} \int_{\Omega_1} \mu_2\big(A^{(n)}(\omega_1)\big)\,\mu_1(d\omega_1) = \sum_{n=1}^{\infty} \mu(A^{(n)}).$$

故 μ 是 $\mathcal{F}_1 \times \mathcal{F}_2$ 上的测度. 同样可证由 (2) 定义的 μ 也是 $\mathcal{F}_1 \times \mathcal{F}_2$ 上的测度. □

对于有限个测度空间的乘积测度空间的存在性，可以用数学归纳法(也可以完全比照两个测度空间的情形)来证明相应的结果. 下面将叙述这些结论而不加证明，读者可以作为练习自己去证明.

11. 定义 设 $A \subset \Omega_1 \times \cdots \times \Omega_n$，$f$ 是 $\Omega_1 \times \cdots \times \Omega_n$ 上的数值函数， $(i_1, \cdots, i_k, j_1, \cdots, j_{n-k})$ 是 $(1, 2, \cdots, n)$ 的一个置换且 $i_1 < \cdots < i_k$，$j_1 < \cdots < j_{n-k}$，$(\omega_{i_1}, \cdots, \omega_{i_k}) \in \Omega_{i_1} \times \cdots \times \Omega_{i_k}$，则称

$$A(\omega_{i_1}, \cdots, \omega_{i_k}) := \Big\{(\omega_{j_1}, \cdots, \omega_{j_{n-k}}):\ (\omega_1, \cdots, \omega_n) \in A\Big\}$$

为 A 在 $(\omega_{i_1}, \cdots, \omega_{i_k})$ 处的 **截集**，称函数 $f_{(\omega_{i_1}, \cdots, \omega_{i_k})}$:

$$f_{(\omega_{i_1}, \cdots, \omega_{i_k})}(\omega_{j_1}, \cdots, \omega_{j_{n-k}}) := f(\omega_1, \cdots, \omega_n),$$

$$(\omega_{j_1}, \cdots, \omega_{j_{n-k}}) \in \Omega_{j_1} \times \cdots \times \Omega_{j_{n-k}},$$

为 f 在 $(\omega_{i_1}, \cdots, \omega_{i_k})$ 处的 **截函数**.

用证明定理 6, 8 的方法, 同样可证明: 若 $A \in \mathcal{F}_1 \times \cdots \times \mathcal{F}_n$ (相应地, f 为 $\mathcal{F}_1 \times \cdots \times \mathcal{F}_n$ 可测函数), 则对一切 $(\omega_{i_1}, \cdots, \omega_{i_k}) \in \Omega_{i_1} \times \cdots \times \Omega_{i_k}$, $A(\omega_{i_1}, \cdots, \omega_{i_k}) \in \mathcal{F}_{j_1} \times \cdots \times \mathcal{F}_{j_{n-k}}$ (相应地 $f_{(\omega_{i_1}, \cdots, \omega_{i_k})}$ 为 $\mathcal{F}_{j_1} \times \cdots \times \mathcal{F}_{j_{n-k}}$ 可测). 即可测集 (函数) 的截口仍然可测, 而类比于定理 9, 10, 则有下列定理 12, 13.

12. 定理　设 f 是非负 $\mathcal{F}_1 \times \cdots \times \mathcal{F}_n$ 可测函数, μ_i 是 $\mathcal{F}_i$, $i = 1, 2, \cdots, n$, 上的 $\sigma-$ 有限测度, 则对一切 $\{i_1, \cdots, i_k\} \subset \{1, 2, \cdots, n\}$,

$$\int_{\Omega_{i_k}} \left(\cdots \left(\int_{\Omega_{i_1}} f(\omega_1, \cdots \omega_n) \, \mu_{i_1}(d\omega_{i_1}) \right) \cdots \right) \mu_{i_k}(d\omega_{i_k})$$

是非负 $\mathcal{F}_{j_1} \times \cdots \times \mathcal{F}_{j_{n-k}}$ 可测函数, 其中 $(i_1, \cdots, i_k, j_1, \cdots, j_{n-k})$ 是 $(1, 2, \cdots, n)$ 的一个置换, 且 $j_1 < \cdots < j_{n-k}$.

13. 定理　设 $(\omega_k, \mathcal{F}_k, \mu_k)$, $k = 1, 2, \cdots, n$, 是 n 个 $\sigma-$ 有限测度空间, 则在乘积 $\sigma-$ 代数 $\mathcal{F}_1 \times \cdots \times \mathcal{F}_n$ 上存在唯一的 $\sigma-$ 有限测度 μ, 满足

$$(1) \qquad \mu(A_1 \times \cdots \times A_n) = \mu_1(A_1) \cdots \mu_n(A_n), \quad A_k \in \mathcal{F}_k, \quad k = 1, 2, \cdots, n.$$

今后称此测度为 μ_k, $k = 1, 2, \cdots, n$, 的 **乘积测度**, 记作 $\mu_1 \times \cdots \times \mu_n$ 或 $\prod_{k=1}^n \mu_k$, 而且它可以用下列累次积分表示:

$$\begin{aligned} & (\mu_1 \times \cdots \times \mu_n)(A) \\ (2) \quad & = \int_{\Omega_{i_n}} \left(\cdots \left(\int_{\Omega_{i_1}} I_A(\omega_1, \cdots \omega_n) \, \mu_{i_1}(d\omega_{i_1}) \right) \cdots \right) \mu_{i_k}(d\omega_{i_n}), \\ & \quad A \in \mathcal{F}_1 \times \cdots \times \mathcal{F}_n, \end{aligned}$$

其中 $(i_1, \cdots, i_n)$ 是 $(1, 2, \cdots, n)$ 的任一置换.

证明　我们先用数学归纳法证明 $\mu_1 \times \cdots \times \mu_n$ 的存在性. 作归纳假设, 则由定理 10 得知存在 $(\mathcal{F}_1 \times \cdots \times \mathcal{F}_{n-1}) \times \mathcal{F}_n$ 上的一

个 $\sigma-$ 有限测度 μ, 满足

$$\mu(A^{(n-1)}\times A_n) = \mu^{(n-1)}(A^{(n-1)})\mu(A_n),$$
$$\forall\ A^{(n-1)} \subset \mathcal{F}_1 \times \cdots \times \mathcal{F}_{n-1}, \quad A_n \in \mathcal{F}_n,$$

于是由归纳假设知 μ 满足

(3) $$\mu((A_1 \times \cdots \times A_{n-1})\times A_n) = \prod_{k=1}^n \mu_k(A_k),$$
$$\forall\ A_k \in \mathcal{F}_k, \quad k = 1, 2, \cdots, n.$$

由定义 4 后面的解释 $(\mathcal{F}_1 \times \cdots \times \mathcal{F}_{n-1}) \times \mathcal{F}_n = \mathcal{F}_1 \times \cdots \times \mathcal{F}_n$, $(A_1 \times \cdots \times A_{n-1}) \times A_n = A_1 \times \cdots \times A_n$. 即知 μ 为 $\mu_1 \times \cdots \times \mu_n$.

由于 $\mathcal{C} := \{A_1 \times \cdots \times A_n:\ A_k \in \mathcal{F}_k,\ k = 1, 2, \cdots, n\}$ 为 $\Omega_1 \times \cdots \times \Omega_n$ 的一个半集代数, $\sigma(\mathcal{C}) = \mathcal{F}_1 \times \cdots \times \mathcal{F}_n$ 及 μ 的 $\sigma-$ 有限性, 应用测度扩张的唯一性知 $\mu_1 \times \cdots \times \mu_n$ 是唯一的.

至于 (2) 式的成立, 可如下得到: (2) 式右边是 $\mathcal{F}_1 \times \cdots \times \mathcal{F}_n$ 上的一个 $\sigma-$ 有限测度, 而且容易计算: 当 $A = A_1 \times \cdots \times A_n$ 时, $I_A = \prod_{k=1}^n I_{A_k}$, 因而 (2) 式右边等于 $\mu_{i_1}(A_{i_1}) \cdots \mu_{i_n}(A_{i_n}) = \mu_1(A_1) \cdots \mu_n(A_n)$, 所以它是满足 (1) 式的一个 $\sigma-$ 有限测度. 故由 $\mu_1 \times \cdots \times \mu_n$ 的唯一性知 (2) 式成立. □

14. 推论 设 μ_k 是 $\mathcal{F}_k$ 上的 $\sigma-$ 有限测度, $k = 1, 2, \cdots, n$, $A \in \mathcal{F}_1 \times \mathcal{F}_2$, 则

$$(\mu_1 \times \mu_2)(A) = 0$$

的充要条件是

(1) $$\mu_2(A(\omega_1)) = 0, \qquad \omega_1 - \text{a.e.}(\mu_1),$$

或

(2) $$\mu_1(A(\omega_2)) = 0, \qquad \omega_2 - \text{a.e.}(\mu_2).$$

证明 因为 $\mu_2(A(\omega_1))$ 非负, 所以由 (10.1) 知 $(\mu_1 \times \mu_2)(A) = 0$ 的充要条件是 (1) 成立, 同样由 (10.2) 可得另一结论. □

与乘积测度在方法上相同的是由转移测度构造乘积空间上的测度. 它是乘积测度的推广, 且在随机过程论中有广泛的应用. 因此在此予以介绍.

15. 定义 设 $(\Omega_i,\mathcal{F}_i)$, $i=1,2$ 是可测空间. 映射 $\lambda:\Omega_1\times\mathcal{F}_2\to[0,\infty]$ 如果满足下列条件, 则称它为 $(\Omega_1,\mathcal{F}_1)$ **到** $(\Omega_2,\mathcal{F}_2)$ **上的转移测度**, 简称为 $\Omega_1\times\mathcal{F}_2$ 上的 **转移测度**:

(i) 对一切 $B\in\mathcal{F}_2$, $\lambda(\cdot,B)$ 是 $\mathcal{F}_1$ - 可测函数;

(ii) 对一切 $\omega\in\Omega_1$, $\lambda(\omega,\cdot)$ 是 $\mathcal{F}_2$ 上的测度.

若存在 $B_{kn}\in\mathcal{F}_k$, $n\in\mathbb{N}$, 两两不交, $\Omega_k=\cup_{n=1}^{\infty}B_{kn}, k=1,2$, 使得

$$\sup_{\omega\in B_{1m}}\lambda(\omega,B_{2n})<\infty,\qquad\forall\, m,n\in\mathbb{N},$$

则称 λ 为 σ- **有限转移测度**.

如果对一切 $\omega\in\Omega_1$, $\lambda(\omega,\Omega_2)=1$, 则称 λ 为 $\Omega_1\times\mathcal{F}_2$ 上的 **转移概率**. 显然转移概率是 σ- 有限转移测度.

例1　设 μ 是 $\mathcal{F}_2$ 上的测度, 令

$$\lambda(\omega,B)=\mu(B),\quad \omega\in\Omega_1,\ B\in\mathcal{F}_2.$$

则 λ 是 $\Omega_1\times\mathcal{F}_2$ 上的转移测度. 因此 $\mathcal{F}_2$ 上的测度可以看作转移测度的特殊情形. 它表示此转移测度与 $\omega\in\Omega_1$ 无关. 而一般的转移测度表示 $\mathcal{F}_2$ 上的与 Ω_1 有某种相关性的测度.

例2　设 $\Omega_1=\Omega_2:=\mathbb{Z}:=\{0,\pm1,\pm2,\cdots\}$, $\mathcal{F}_1=\mathcal{F}_2:=\{B:B\subset\mathbb{Z}\}$,

$$p_{ij}:=\begin{cases}\frac{1}{2}, & \text{当 } |j-i|=1,\\ 0, & \text{当 } |j-i|\neq1,\ i,j\in\mathbb{Z};\end{cases}$$

$$\lambda(i,B):=\sum_{j\in B}p_{ij},\quad i\in\mathbb{Z},\ B\subset\mathbb{Z}.$$

这是一个转移概率, 通常称为**简单对称随机游动**的转移概率.

例3 设

$$\mathbf{P}=\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix},$$

其中 $p_{ij}\geq 0$. 每行的和为 1. 即 $\sum_j p_{ij}=1, i=1,\cdots,n$. 设 $\Omega_1=\Omega_2:=\{1,\cdots,n\}, \mathcal{F}_1=\mathcal{F}_2:=\{B:\ B\subset\Omega_2\}$. 令

$$\lambda(i,B):=\sum_{j\in B}p_{ij},\qquad i\in\Omega_1,\ B\in\mathcal{F}_2.$$

则 λ 是 $\Omega_1\times\mathcal{F}_2$ 上的转移概率. 称 $\mathbf{P}$ 为它的 **转移概率矩阵**. 这里的 $\mathbf{P}$ 也是有限状态马链的转移概率矩阵.

例4 设

$$\mathbf{P}=\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} & \cdots \\ p_{21} & p_{22} & \cdots & p_{2n} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

其中 $p_{ij}\geq 0$. 每行的和为 1. 即 $\sum_j p_{ij}=1, i\in\mathbb{Z}$. 设 $\Omega_1=\Omega_2:=\mathbb{Z}, \mathcal{F}_1=\mathcal{F}_2:=\{B:\ B\subset\Omega_1\}$. 令

$$\lambda(i,B):=\sum_{j\in B}p_{ij},\qquad i\in\Omega_1,\ B\in\mathcal{F}_2.$$

则 λ 是 $\Omega_1\times\mathcal{F}_2$ 上的转移概率. 称 $\mathbf{P}$ 为它的转移概率矩阵. 这里的 $\mathbf{P}$ 也是可数状态马链的转移概率矩阵，显然例 (2). 例 (3) 都是本例的特别情形. 马链的转移概率常用来表示一个系统的下述运动规律：设系统状态空间为 $\mathbb{Z}_+$, 在每一单位时间进行一次状态的随机改变，而且改变的概率与出发的时刻无关. 此时可以用 p_{ij} 表示从状态 i 出发在一个单位时间转移到状态 j 的概率.

例5 设 $\Omega_i=\mathbb{R}^3$, $\mathcal{F}_i=\mathcal{B}^3$, $i=1,2$. 对一切 $t\geq 0$, 令

$$p(t;x,A):=(2\pi t)^{-3/2}\int_A e^{-\frac{|y-x|^2}{2t}}\,dy,\quad x\in\mathbb{R}^3,\ A\in\mathcal{B}^3.$$

则对给定的 $t\geq 0$,$p(t;\cdot,\cdot)$ 是 $\mathbb{R}\times\mathcal{B}$ 上的转移概率. $p(\cdot;\cdot,\cdot)$ 即一维布朗运动的转移概率. 它的直观背景如下：考察水中一悬浮粒子运动的规律, 由于众多水分子的撞击, 粒子的运动完全是随机的. 它的概率规律可以上述转移概率描述. 即粒子从 x 处出发, 经过时间 t, 进入 A 的概率可以 $p(t;x,A)$ 表示 (假设已选定坐标).

由一个测度和一个转移测度出发, 用构造乘积测度的方法可以唯一地构造乘积空间上的测度. 为此先给出下列的

16. 定理 给定 $(\Omega_i,\mathcal{F}_i)$, $i=1,2$. 若 f 是非负 $\mathcal{F}_1\times\mathcal{F}_2$ 可测函数, λ 是 $\Omega_1\times\mathcal{F}_2$ 上的 $\sigma-$ 有限转移测度, 则 $f(\omega_1,\cdot)$ 对 $\lambda(\omega_1,\cdot)$ 的积分

$$\int_{\Omega_2} f(\cdot,\omega_2)\lambda(\cdot,d\omega_2)$$

是非负 $\mathcal{F}_1$ 可测函数.

证法与定理 9 完全相同.

17. 定理 给定 $(\Omega_i,\mathcal{F}_i)$, $i=1,2$. 若 λ_1, λ_2 分别是 $\mathcal{F}_1$ 及 $\Omega_1\times\mathcal{F}_2$ 上的 $\sigma-$ 有限测度, 则由

$$\begin{aligned}\lambda(B)&:=\int_{\Omega_1}\int_{\Omega_2} I_B(\omega_1,\omega_2)\lambda_2(\omega_1,d\omega_2)\lambda_1(d\omega_1)\\&=\int_{\Omega_1}\lambda_2(\omega_1,B(\omega_1))\lambda_1(d\omega_1),\qquad B\in\mathcal{F}_1\times\mathcal{F}_2,\end{aligned}$$

定义的 λ 是 $\mathcal{F}_1\times\mathcal{F}_2$ 上的 $\sigma-$ 有限测度.

证法与定理 10 相同, 略.

与定理 13 相当的有关转移测度的结果是下列定理.

18. **定理** 给定 n 个可测空间 $(\Omega_i, \mathcal{F}_i)$, $i=1,\cdots,n$. 设 $\Omega^{(k)} := \prod_{i=1}^k \Omega_i$, $\mathcal{F}^{(k)} := \prod_{i=1}^k \mathcal{F}_i$, $k=1,\cdots,n$. 若 λ_1 是 $\mathcal{F}_1$ 上的 $\sigma-$ 有限测度, λ_k 是 $\Omega^{(k-1)} \times \mathcal{F}_k$ 上的 $\sigma-$ 有限转移测度, $k=2,\cdots,n$, 则由

$$\lambda^{(n)}(B) := \int_{\Omega_1} \cdots \int_{\Omega_n} I_B(\omega_1, \cdots, \omega_n)$$
$$\lambda_n(\omega_1, \cdots, \omega_{(n-1)}; d\omega_n) \cdots \lambda_1(d\omega_1),$$

$B \in \mathcal{F}^{(n)}$, 定义的 $\lambda^{(n)}$ 是 $\mathcal{F}^{(n)}$ 上的 $\sigma-$ 有限测度.

证明与定理 16 的方法相同, 略.

最后, 我们介绍乘积测度存在定理对 L-S 测度及概率论的应用.

19. **推论** 设 μ_k 是 $\mathcal{B}^{m_k}$ 上的 L-S 测度, μ_k 由分布函数 F_k 决定, $k=1,2,\cdots,n$, 则 $\mu_1 \times \cdots \times \mu_n$ 由分布函数

(1) $\quad F(x^{(1)}, \cdots, x^{(n)}) := \prod_{k=1}^n F_k(x^{(k)}), \qquad x^{(k)} \in \mathbb{R}^{m_k},$

决定, 且若用 μ_F 表示由 F 决定的测度, 则对一切 $B_k \in \mathcal{B}^{m_k}$, $k=1,2,\cdots,n$, 有

(2) $$\mu_F(B_1 \times \cdots \times B_n) = \prod_{k=1}^n \mu_{F_k}(B_k).$$

特别, 用 $\lambda^{(d)}$ 表示 $\mathcal{B}^d (d \in \mathbb{N})$ 上的 Lebesgue 测度, 则

(3) $$\lambda^{(n)} = \underbrace{\lambda^{(1)} \times \cdots \times \lambda^{(1)}}_{n} = \lambda^{(k)} \times \lambda^{(n-k)}, \qquad 1 < k < n.$$

证明 设 $a^{(k)}, b^{(k)} \in \mathbb{R}^{m_k}$, $a^{(k)} \leq b^{(k)}$, $k=1,2,\cdots,n$, 则由假设知

(4) $$\mu_k((a^{(k)}, b^{(k)}]) = \Delta_{b^{(k)}, a^{(k)}} F_k.$$

令

$$a := (a^{(1)}, \cdots, a^{(n)}) \in \mathbb{R}^{m_1 + \cdots + m_n},$$

$$b := (b^{(1)}, \cdots, b^{(n)}) \in \mathbb{R}^{m_1+\cdots+m_n}.$$

由 (1) 容易算得

(5) $$\Delta_{b,a}F = \prod_{k=1}^{n} \Delta_{b^{(k)},a^{(k)}} F_k.$$

而由 F 决定的测度 $\mu_F((a,b]) = \Delta_{b,a}F$, 所以由 (4)、(5) 得

$$(\mu_1 \times \cdots \times \mu_n)((a,b]) = \prod_{k=1}^{n} \mu_k((a^{(k)}, b^{(k)}]) = \mu_F((a,b]).$$

再由 $\mathcal{S} := \{(a,b]: a,b \in \mathbb{R}^{m_1+\cdots+m_n}, a \leq b\}$ 是半集代数及 $\mu_1 \times \cdots \times \mu_n$、$\mu_F$ 的 $\sigma-$ 有限性及测度扩张定理知 $\mu_1 \times \cdots \times \mu_k = \mu_F$. 这就是第一个结论. 再设 $B_k \in \mathcal{B}^{m_k}$, $k = 1, 2, \cdots, n$, 则由第一结论及假设知

$$\begin{aligned}\mu_F(B_1 \times \cdots \times B_n) &= (\mu_1 \times \cdots \times \mu_n)(B_1 \times \cdots \times B_n) \\ &= \prod_{k=1}^{n} \mu_k(B_k) = \prod_{k=1}^{n} \mu_{F_k}(B_k).\end{aligned}$$

故 (2) 获证. 至于 (3), 显然是第一结论的推论. □

20. 推论 设 X_k 是 $(\Omega_k, \mathcal{F}_k, \mathbf{P}_k)$ 上的 m_k 维 r.v., $k = 1, 2, \cdots, n$, 而 $(\Omega, \mathcal{F}, \mathbf{P})$ 是上述 n 个概率空间的乘积概率空间, 在 $(\Omega, \mathcal{F}, \mathbf{P})$ 上定义 r.v. $Y_1, \cdots, Y_n$ 如下:

(1) $$Y_k(\omega_1, \cdots, \omega_n) := X_k(\omega_k), \quad \omega_k \in \Omega_k,$$
$$k = 1, 2, \cdots, n, (\omega_1, \cdots, \omega_n) \in \Omega,$$

则 $Y_1, \cdots, Y_n$ 是独立 r.v., 且 Y_k 与 X_k 同分布, $k = 1, 2, \cdots, n$.

由此即得: 对于任意给定的 n 个 (多元) 概率分布函数 F_k, $k = 1, 2, \cdots, n$, 总存在 n 个独立 r.v. $Y_1, \cdots, Y_n$, 使得 Y_k 的分布函数是 $F_k k = 1, 2, \cdots, n$.

证明 由于对一切 $x^{(k)} \in \mathbb{R}^{m_k}$, $k = 1,2,\cdots,n$,

$$\{(\omega_1,\cdots,\omega_n):\ Y_k(\omega_1,\cdots,\omega_n) \le x^{(k)}, k = 1,2,\cdots,n\}$$
$$= \{(\omega_1,\cdots,\omega_n):\ X_k(\omega_k) \le x^{(k)},\quad k = 1,2,\cdots,n\}$$
$$= \{\omega_1 \in \Omega_1:\ X_1(\omega_1) \le x^{(1)}\} \times \cdots$$
$$\times \{\omega_n \in \Omega_n:\ X_n(\omega_n) \le x^{(n)}\}$$

于是, 因为 $\mathbf{P} = \mathbf{P}_1 \times \cdots \times \mathbf{P}_n$,

$$\mathbf{P}(Y_1 \le x^{(1)},\cdots,Y_n \le x^{(n)})$$
$$= (\mathbf{P}_1 \times \cdots \times \mathbf{P}_n)\left(\{X_1 \le x^{(1)}\} \times \cdots \times \{X_n \le x^{(n)}\}\right)$$
$$= \prod_{k=1}^n \mathbf{P}_k(X_k \le x^{(k)}) = \prod_{k=1}^n \mathbf{P}(Y_k \le x^{(k)}).$$

故 $Y_1,\cdots,Y_n$ 独立, 且 $\mathbf{P}_k(X_k \le x^{(k)}) = \mathbf{P}(Y_k \le x^{(k)})$.

设 F_k 是 m_k 元分布函数. 则由 $X_k(x_k) := x_k \in \mathbb{R}^{m_k}$ 定义的函数 X_k 是 $(\mathbb{R}^{m_k},\mathcal{B}^{m_k},\mu_{F_k})$ 上的 r.v., 且它的概率分布函数 F_k. 故由已证结论即得后一结论. □

应用推论 19 可以重新证明独立事件类扩张定理, 这就是下列的

21. 推论 设 m_k 元 r.v. X_k 的概率分布函数为 F_k, $k = 1,2,\cdots,n$, 而且独立. 则对一切 $B_k \in \mathcal{B}^{m_k}$, $k = 1,2,\cdots,n$

(1) $\quad \mathbf{P}(X_1 \in B_1,\cdots,X_n \in B_n) = \mathbf{P}(X_1 \in B_1)\cdots\mathbf{P}(X_n \in B_n).$

因而, 若 f_k 为 $\mathbb{R}^{m_k}$ 上的 Borel 可测函数, 则 $f_k(X_k)$, $k = 1,2,\cdots,n$, 为独立 r.v..

证明 对一切 $x^{(k)} \in \mathbb{R}^{m_k}$, $k = 1,2,\cdots,n$, $(X_1,\cdots,X_n) = X$ 是概率分布函数

$$F(x^{(1)},\cdots,x^{(n)}) = \prod_{k=1}^n F_k(x^{(k)}).$$

于是由推论 19(2) 知

$$\mathbf{P}(X_1 \in B_1, \cdots, X_n \in B_n) = \mu_F(B_1 \times \cdots \times B_n)$$
$$= \prod_{k=1}^n \mu_{F_k}(B_k) = \prod_{k=1}^n \mathbf{P}(X_k \in B_k).$$

故 (1) 获证，而

$$\mathbf{P}(f_1(X_1) \leq x_1, \cdots, f_n(X_n) \leq x_n)$$
$$= \mathbf{P}(X_1 \in f^{-1}((-\infty, x_1]), \cdots, X_n \in f^{-1}((-\infty, x_n]))$$
$$= \prod_{k=1}^n \mathbf{P}(X_k \in f^{-1}((-\infty, x_k])) = \prod_{k=1}^n \mathbf{P}(f_k(X_k) \leq x_k).$$

故 $f_k(X_k), k = 1, 2, \cdots, n$ 独立. □

习题

1. 证明推论 11, 12.

2. 设 Ω 是一不可数集, $\mathcal{F}$ 是包含 Ω 中一切单点集的最小 $\sigma-$ 代数, 则 $\Omega \times \Omega$ 的对角线 $\Delta := \{(\omega, \omega) : \omega \in \Omega\} \notin \mathcal{F} \times \mathcal{F}$, 但对任意 $\omega_i \in \Omega, i = 1, 2$

$$\Delta_{\omega_1} := \{\omega_2 : (\omega_1, \omega_2) \in \Delta\} \in \mathcal{F},$$
$$\Delta_{\omega_2} := \{\omega_1 : (\omega_1, \omega_2) \in \Delta\} \in \mathcal{F}.$$

这个例子说明了什么？

3. 试问： $\overline{\mathcal{F}}_1 \times \overline{\mathcal{F}}_2 = \overline{\mathcal{F}_1 \times \mathcal{F}_2}$ 吗？其中 $\overline{\mathcal{F}}_i, i = 1, 2$ 表示 $\mathcal{F}_i$ 对 μ_i 的完全化, $\overline{\mathcal{F}_1 \times \mathcal{F}_2}$ 表示 $\mathcal{F}_1 \times \mathcal{F}_2$ 对 $\mu_1 \times \mu_2$ 的完全化, 这个问题对 Lebesgue 可测集说明了什么？

4. 设 X_1, X_2 是 n 维独立 r.v. , $\mathbf{P}_i, F_i$, 分别是 X_i 的概率分布测度和分布函数, $i = 1, 2$.

1) 试用乘积概率定理证明 $X_1 + X_2$ 的概率分布测度和分布函

数分别为由

$$\mathbf{P}_1 * \mathbf{P}_2(B) := \int_{\mathbb{R}^n} \mathbf{P}_1(B-y)\mathbf{P}_2(dy), \quad B \in \mathcal{B}^n,$$

$$F_1 * F_2(x) := \int_{\mathbb{R}^n} F_1(x-y)\, dF_2(y), \quad x \in \mathbb{R}^n,$$

定义的 $\mathbf{P}_1 * \mathbf{P}_2$, $F_1 * F_2$; 它们分别称为 $\mathbf{P}_1, \mathbf{P}_2$ 及 F_1, F_2 的 **卷积**.

2) 若 X_i, $i = 1, 2$, 还具有分布密度 p_i. 则由

$$p_1 * p_2(x) := \int_{\mathbb{R}^n} p_1(x-y)p_2(y)\, dy, \quad x \in \mathbb{R}^n$$

定义的 $p_1 * p_2$ 是 $X_1 + X_2$ 的分布密度. $p_1 * p_2$ 也称为 p_1, p_2 的卷积.

3) 试证: 一切概率分布测度 (相应地: 分布函数) 对卷积运算作成一个可交换半群.

5. 设 $(\Omega, \mathcal{F}, \mathbf{P})$ 是概率空间, $f(t,\omega)$ 作为 $(t,\omega) \in \mathbb{R} \times \Omega$, 的函数是 $\mathcal{B} \times \mathcal{F}$ 可测的, 若对一切 $t \in \mathbb{R}$, $\mathbf{P}\left(\left\{\omega \in \Omega:\ f(t,\omega) = \infty\right\}\right) = 0$, 试证

$$\lambda\left(\left\{t \in \mathbb{R}:\ f(t,\omega) = \infty\right\}\right) = 0, \qquad \text{a.e. } \omega(\mathbf{P}),$$

其中 λ 表示 $\mathbb{R}$ 上的 Lebesgue 测度. (提示: 令 $A := \left\{(t,\omega):\ f(t,\omega) = \infty\right\}$, 考虑 $(\lambda \times \mathbf{P})(A)$)

6. 设 f 是 $(\Omega_1 \times \cdots \times \Omega_n, \mathcal{F}_1 \times \cdots \times \mathcal{F}_n)$ 上的 (实或复) 数值可测函数, 设 $(i_1, \cdots, i_k, j_1, \cdots, j_{n-k})$ 是 $(1, 2, \cdots, n)$ 的一个置换, $j_1 < \cdots < j_{n-k}$. 试证: 对任何 $(\omega_{i_1}, \cdots, \omega_{i_k}) \in \Omega_{i_1} \times \cdots \times \Omega_{i_k}$, f 在 $(\omega_{i_1}, \cdots, \omega_{i_k})$ 截函数 $f_{(\omega_{i_1}, \cdots, \omega_{i_k})}$ 是 $\mathcal{F}_{j_1} \times \cdots \times \mathcal{F}_{j_{n-k}}$ 可测的.

7. 试证定理 16.

8. 试证定理 17 及 18.

9. 设 $f(x,y)$ 是 $[0,1] \times [0,1]$ 上的非负有界可测函数, 对一切

$B \in \mathcal{B}[0,1]$. 令

$$\lambda(x,B) := \int_B f(x,y)\,dy.$$

其中 dy 表示对 *Lebesgue* 测度的积分. 则 λ 是 $[0,1] \times \mathcal{B}[0,1]$ 上的转移测度.

10. 设 $(\Omega_i, \mathcal{F}_i)$, $i = 1,2$ 是可测空间, λ_1 是 $\mathcal{F}_1$ 上的 $\sigma-$ 有限测度, λ_2 是 $\Omega_1 \times \mathcal{F}_2$ 上的 $\sigma-$ 有限转移测度,

$$\nu(B) := \int_{\Omega_1}\int_{\Omega_2} I_B(\omega_1,\omega_2)\lambda_2(\omega_1, d\omega_2)\lambda_1(d\omega_1), \quad B \in \mathcal{F}_1 \times \mathcal{F}_2.$$

$A \in \mathcal{F}_1 \times \mathcal{F}_2$, 则 $\nu(A) = 0$ 的充分与必要条件是存在一个 λ_1 零测集 N, 使对一切 $\omega_1 \in N^c$,

$$\lambda_2(\omega_1, A(\omega_1)) = 0.$$

11. 设 $(\Omega_i, \mathcal{F}_i)$, $i = 1,2,3$ 是可测空间, λ 是 $\Omega_2 \times \mathcal{F}_3$ 上的 $\sigma-$ 有限转移测度, f 是 $\mathcal{F}_1 \times \mathcal{F}_3$ 可测函数, 若积分

$$g(\omega_1,\omega_2) := \int_{\Omega_3} f(\omega_1,\omega_3)\lambda(\omega_2, d\omega_3), \quad \omega_i \in \Omega_i, i = 1,2,$$

对一切 $\omega_i, i = 1,2$ 存在, 则 g 是 $\mathcal{F}_1 \times \mathcal{F}_2$ 可测的.

12. 设 $(\Omega_i, \mathcal{F}_i)$, $i = 1,2,3,4$, 是可测空间, λ_1, λ_2 分别是 $\Omega_1 \times \mathcal{F}_2$, $\Omega_2 \times \mathcal{F}_3$ 上的转移概率. 则由

$$\lambda_1 \circ \lambda_2(\omega, B) := \int \lambda_2(\omega_2, B)\lambda_1(\omega, d\omega_2), \quad \omega \in \Omega_1, B \in \mathcal{F}_3$$

定义的 $\lambda_1 \circ \lambda_2$ 是 $\Omega_1 \times \mathcal{F}_3$ 上的转移概率. 若 λ_3 是 $\Omega_3 \times \mathcal{F}_4$ 上的转移概率, 则

$$(\lambda_1 \circ \lambda_2) \circ \lambda_3 = \lambda_1 \circ (\lambda_2 \circ \lambda_3).$$

13. 设 λ_i $i = 1,2$ 是由转移概率矩阵 $\mathbf{P}_i$ 确定的 $\mathbb{N} \times \mathcal{N}$, $\mathcal{N} := \{A: A \subset \mathbb{N}\}$ 上的转移概率. 则 $\lambda_1 \circ \lambda_2$ 是由 $\mathbf{P}_1 \cdot \mathbf{P}_2$ (矩阵 $\mathbf{P}_1$, $\mathbf{P}_2$ 的乘积) 确定的.

14. 设 $p(t;x,A)$, $(t,x,A)\in[0,\infty)\times\mathbb{R}^3\times\mathcal{B}^3$, 是第 11 目所定义的. 试证: 对一切 $t,s\in[0,\infty)$, $x\in\mathbb{R}^3$, $A\in\mathcal{B}^3$,

$$p(t+s;x,A)=\int_{\mathbb{R}^3}p(t;x,dy)p(s;y,A).$$

15. 设 $(\Omega,\mathcal{F})$ 为可测空间, $\mathcal{C}$ 为 $\mathcal{F}$ 的子 $\sigma-$ 代数, 称 $\pi(x,A)$, $x\in\Omega$, $A\in\mathcal{F}$, 为一概率核, 若它对每一 $A\in\mathcal{F}$, $\pi(\cdot,A)$ 为 $\mathcal{C}$ 可测函数, 对每一 $x\in\Omega$, $\pi(x,\cdot)$ 是 $\mathcal{F}$ 上的概率. 设 ν 是 $\mathcal{F}$ 上任一概率, 试证

$$\nu\pi(\cdot):=\int\pi(x,\cdot)\nu(dx)$$

为 $\mathcal{F}$ 上的概率测度.

§2. Fubini 定理及其应用

这一节讨论乘积测度空间 $(\Omega_1\times\cdots\times\Omega_n,\mathcal{F}_1\times\cdots\times\mathcal{F}_n,\mu_1\times\cdots\times\mu_n)$ 上的积分 (重积分) 与各因子测度空间 $(\Omega_k;\mathcal{F}_k,\mu_k)$, $k=1,2,\cdots,n$, 上的积分的关系, 我们将证明在一定条件下 "重" 积分可以化为对 $\mu_1,\cdots,\mu_n$ 的累次积分进行积分, 通常称为 Fubini 定理, 并介绍它的几个应用.

我们首先讨论 $n=2$ 的情形, 象 §1 一样, 它是 n 为有限时最本质的情形.

1. 定理 设 $(\Omega_k,\mathcal{F}_k,\mu_k)$, $k=1,2$ 是两个 $\sigma-$ 有限测度空间, 若 f 是非负 $\mathcal{F}_1\times\mathcal{F}_2$ 可测函数, 则

$$\begin{aligned}\int_{\Omega_1\times\Omega_2}&f\,d(\mu_1\times\mu_2)\\&=\int_{\Omega_1}\left(\int_{\Omega_2}f(\omega_1,\omega_2)\,\mu_2(d\omega_2)\right)\mu_1(d\omega_1)\\&=\int_{\Omega_2}\left(\int_{\Omega_1}f(\omega_1,\omega_2)\,\mu_1(d\omega_1)\right)\mu_2(d\omega_2)\end{aligned}$$

证明　由于方法相同，只证前一等式.

由定理 1.9 知 $\int_{\Omega_2} f(\cdot,\omega_2)\,\mu_2(d\omega_2)$ 非负 $\mathcal{F}_1$ 可测，所以 (1) 中第二个积分存在. 其次，当 $A \in \mathcal{F}_1 \times \mathcal{F}_2$ 时，由定理 1.10 知

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} I_A\, d(\mu_1\times\mu_2) &:= (\mu_1\times\mu_2)(A)\\ &= \int_{\Omega_1} \mu_2(A(\omega_1))\,\mu_1(d\omega_1)\\ &= \int_{\Omega_1}\left(\int_{\Omega_2} I_{A(\omega_1)}(\omega_2)\,\mu_2(d\omega_2)\right)\mu_1(d\omega_1)\\ &= \int_{\Omega_1}\left(\int_{\Omega_2} I_A(\omega_1,\omega_2)\,\mu_2(d\omega_2)\right)\mu_1(d\omega_1)\end{aligned}$$

即结论的第一式对 $f = I_A$, $A \in \mathcal{F}_1\times\mathcal{F}_2$, 成立，由积分的线性性质知结论对非负 $\mathcal{F}_1\times\mathcal{F}_2$ 简单函数成立，由单调收敛定理及非负可测函数是某一不降非负可测简单函数序列的极限知结论的第一式对非负可测函数成立，故定理获证.

2. 定理　(Fubini 定理) 设 f 是 $\mathcal{F}_1\times\mathcal{F}_2$ 可测函数且 $\int f\,d(\mu_1\times\mu_2)$ 存在，则

1) $g(\omega_1) := \int_{\Omega_2} f(\omega_1,\omega_2)\,\mu_2(d\omega_2)$　$\omega_1 - a.e.(\mu_1)$ 存在且 $\mathcal{F}_1$ 可测；

$h(\omega_2) := \int_{\Omega_1} f(\omega_1,\omega_2)\,\mu_1(d\omega_1)$　$\omega_2 - a.e.(\mu_2)$ 存在且 $\mathcal{F}_2$ 可测；

2) $\int_{\Omega_1} g\,d\mu_1$, $\int_{\Omega_2} h\,d\mu_2$ 都存在且

$$\int_{\Omega-1\times\Omega_2} f\,d(\mu_1\times\mu_2) = \int_{\Omega_1} g\,d\mu_1 = \int_{\Omega_2} h\,d\mu_2. \tag{1}$$

3) 若 f 对 $\mu_1\times\mu_2$ 可积，则 g、h 分别对 μ_1,μ_2 可积.

注　定理 2 实际包括定理 1.

证明　由于方法相同，所以只证关于 g 的论断. 又由于复值函数可分实虚两部考虑，所以只需考虑 f 为实函数的情形.

1) 由于 $\int_{\Omega_1\times\Omega_2} f\,d(\mu_1\times\mu_2)$ 存在，所以 $\int_{\Omega_1\times\Omega_2} f^+\,d(\mu_1\times\mu_2)$ 有限，或 $\int_{\Omega_1\times\Omega_2} f^-\,d(\mu_1\times\mu_2)$ 有限，不妨设后者有限．由定理 1 知

$$\int_{\Omega_1\times\Omega_2} f^{\pm}\,d(\mu_1\times\mu_2) = \int_{\Omega_1}\left(\int_{\Omega_2} f^{\pm}(\omega_1,\omega_2)\,\mu_2(d\omega_2)\right)\mu_1(d\omega_1).$$

因而 $\int_{\Omega_2} f^-(\omega_1,\omega_2)\,\mu_2(d\omega_2)$ $\omega_1 - a.e.(\mu_1)$ 有限且作为 ω_1 的函数，对 μ_1 是可积的，又由定理 1.9 知 $\int_{\Omega_2} f^{\pm}(\omega_1,\omega_2)\,\mu_2(d\omega_2)$ 作为 ω_1 的函数 $\mathcal{F}_1-$ 可测，所以

$$g(\omega_1) = \int_{\Omega_2} f(\omega_1,\omega_2)\,\mu_2(d\omega_2)$$
$$= \int_{\Omega_2} f^+(\omega_1,\omega_2)\,\mu_2(d\omega_2) - \int_{\Omega_2} f^-(\omega_1,\omega_2)\,\mu_2(d\omega_2)$$

ω_1-a.e. 有意义 (即存在) 且 $\mathcal{F}_1-$ 可测，故 1) 获证．

2) 由积分的线性性质及 $\int_{\Omega_2} f^-(\omega_1,\omega_2)\,\mu_2(d\omega_2)$ ω_1- a.e. (μ_1) 有限且对 μ_1 可积，$\int_{\Omega_1\times\Omega_2} f\,d(\mu_1\times\mu_2)$ 存在．所以

$$\int_{\Omega_1\times\Omega_2} f\,d(\mu_1\times\mu_2)$$
$$= \int_{\Omega_1\times\Omega_2} f^+\,d(\mu_1\times\mu_2) - \int_{\Omega_1\times\Omega_2} f^-\,d(\mu_1\times\mu_2)$$
$$= \int_{\Omega_1}\left(\int_{\Omega_2} f^+(\omega_1,\omega_2)\,\mu_2(d\omega_2)\right)\mu_1(d\omega_1)$$
$$- \int_{\Omega_1}\left(\int_{\Omega_2} f^-(\omega_1,\omega_2)\,\mu_2(d\omega_2)\right)\mu_1(d\omega_1)$$
$$= \int_{\Omega_1}\left(\int_{\Omega_2} f^+(\omega_1,\omega_2)\,\mu_2(d\omega_2)\right.$$

$$-\int_{\Omega_2} f^-(\omega_1,\omega_2)\,\mu_2(d\omega_2)\Big)\mu_1(d\omega_1)$$
$$=\int_{\Omega_1} g(\omega_1)\,\mu_1(d\omega_1)$$

3) 若 f 对 $\mu_1\times\mu_2$ 可积，则由 2) 知 $\int_{\Omega_1} f^{\pm}(\cdot,\omega_2)\,\mu_2(d\omega_1)$ 对 μ_1 可积，故 g 对 μ_1 可积. □

3. 推论 设 μ_k 是 $\mathcal{F}_k$ 上的 $\sigma-$ 有限测度， $k=1,2$, f 为 $\mathcal{F}_1\times\mathcal{F}_2$ 可测，若

$$\int_{\Omega_2}\left(\int_{\Omega_1}|f(\omega_1,\omega_2)|\,\mu_2(d\omega_2)\right)\mu_1(d\omega_1)<\infty, \tag{1}$$

或

$$\int_{\Omega_1}\left(\int_{\Omega_2}|f(\omega_1,\omega_2)|\,\mu_1(d\omega_1)\right)\mu_2(d\omega_2)<\infty. \tag{2}$$

则， f 对 $\mu_1\times\mu_2$ 可积，因而

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} f\,d(\mu_1\times\mu_2)&=\int_{\Omega_1}\left(\int_{\Omega_2} f(\omega_1,\omega_2)\,\mu_2(d\omega_2)\right)\mu_1(d\omega_1)\\&=\int_{\Omega_2}\left(\int_{\Omega_1} f(\omega_1,\omega_2)\,\mu_1(d\omega_1)\right)\mu_2(d\omega_2)\end{aligned} \tag{3}$$

证明 $|f|$ 为非负 $\mathcal{F}_1\times\mathcal{F}_2$ 可测，所以由定理 1 知当 (3) 式中的 f 换成 $|f|$ 时成立，因而当 (1), (2) 有一成立时， f 对 $\mu_1\times\mu_2$ 可积，由定理 2. 2) 即得结论.

可以将定理 2、推论 3 推广到积分区域是 $G\in\mathcal{F}_1\times\mathcal{F}_2$ 的情形，这只要将定理 2、推论 3 应用到 $f\cdot I_G$ 上即得

4. 推论 设 μ_k, $k=1,2$, 是 $\sigma-$ 有限测度空间， f 为 $\mathcal{F}_1\times\mathcal{F}_2$ 可测函数.

1) 若 $A_k \in \mathscr{F}_k$, $k = 1, 2$, $fI_{A_1 \times A_2}$ 满足定理 2 或推论 3 中 f 所满足的条件，则

$$\begin{aligned}\int_{A_1 \times A_2} f\, d(\mu_1 \times \mu_2) &= \int_{A_2} \left(\int_{A_1} f(\omega_1, \omega_2)\, \mu_1(d\omega_1) \right) \mu_2(d\omega_2) \\ &= \int_{A_1} \left(\int_{A_2} f(\omega_1, \omega_2)\, \mu_2(d\omega_2) \right) \mu_1(d\omega_1).\end{aligned}$$

2) 若 $G \in \mathscr{F}_1 \times \mathscr{F}_2$, fI_G 满足定理 2 或推论 3 中 f 所满足的条件，则

$$\begin{aligned}\int_G f\, d(\mu_1 \times \mu_2) &= \int_{\Omega_1} \left(\int_{G(\omega_1)} f(\omega_1, \omega_2)\, \mu_2(d\omega_2) \right) \mu_1(d\omega_1) \\ &= \int_{\Omega_2} \left(\int_{G(\omega_2)} f(\omega_1, \omega_2)\, \mu_1(d\omega_1) \right) \mu_2(d\omega_2).\end{aligned}$$

和 §1 一样，上述结论可以毫无困难地推广到有限维 (即有限个测度空间的) 乘积空间上去. 现在我们将它们叙述如下:

5. 定理 设 $(\Omega_k, \mathscr{F}_k, \mu_k)$, $k = 1, 2, \cdots, n$, 是 $\sigma-$ 有限测度空间， f 是 $\mathscr{F}_1 \times \cdots \times \mathscr{F}_n$ 可测函数, 且 f 对 $\mu_1 \times \cdots \times \mu_n$ 的积分存在， $(i_1, \cdots, i_k, j_1, \cdots, j_{n-k})$ 是 $(1, \cdots, n)$ 的一个置换, $j_1 < \cdots < j_{n-k}$, 则

$$g(\omega_{j_1}, \cdots, \omega_{j_{n-k}}) := \int_{\Omega_{i_k}} \left(\cdots \left(\int_{\Omega_{i_1}} f(\omega_1, \cdots, \omega_n) \mu_{i_1}(d\omega_{i_1}) \right) \cdots \right) \mu_{i_k}(d\omega_{i_k}), \tag{1}$$

除去一个 $A_{j_1} \times \cdots \times A_{j_{n-k}}$ 可测且其 $\mu_{j_1} \times \cdots \times \mu_{j_{n-k}}$ 测度为零的

集外，存在且可测.

$$\begin{aligned}(2)\qquad &\int_{\Omega_1\times\cdots\times\Omega_n} f\,d(\mu_1\times\cdots\times\mu_n)\\ &=\int_{\Omega_{j_1}\times\cdots\times\Omega_{j_{n-k}}} g\,d(\mu_{j_1}\times\cdots\times\mu_{j_{n-k}}).\end{aligned}$$

特别，取 $k=n-1$, $j_1=i_n$, 即得

$$\begin{aligned}(3)\qquad &\int_{\Omega_1\times\cdots\times\Omega_n} f\,d(\mu_1\times\cdots\times\mu_n)\\ &=\int_{\Omega_{i_n}}\left(\cdots\left(\int_{\Omega_{i_1}} f(\omega_1,\cdots,\omega_n)\,\mu_{i_1}(d\omega_{i_1})\right)\cdots\right)\mu_{i_n}(d\omega_{i_n}).\end{aligned}$$

若 f 对 $\mu_1\times\cdots\times\mu_n$ 可积，则 g 对 $\mu_{j_1}\times\cdots\times\mu_{j_{n-k}}$ 亦可积.

下面的推论是很有用的.

6. 推论　设 f 是 $\mathcal{F}_1\times\cdots\times\mathcal{F}_n$ 可测函数，若对 $(1,2,\cdots,n)$ 的某一置换 $(i_1,\cdots,i_n)$, 有

$$\int_{\Omega_{i_n}}\left(\cdots\left(\int_{\Omega_{i_1}}|f(\omega_1,\cdots,\omega_n)|\mu_{i_1}(d\omega_{i_1})\right)\cdots\right)\mu_{i_n}(d\omega_{i_n})<\infty,$$

则 f 对 $\mu_1\times\cdots\times\mu_n$ 可积，因而 (5.2) 成立.

相应于由转移测度构造的乘积空间上的测度（定理 1.18）也有类似的结果．即

7. 定理　在定理 1.18 的假设下，并采用该处的记号．又设 f 是 $(\Omega^{(n)},\mathcal{F}^{(n)})$ 上的可测函数．则 f 对 $\lambda^{(n)}$ 的积分存在的充分与必要条件是

$$\int_{\Omega_1}\cdots\int_{\Omega_n} f^{\pm}(\omega_1,\cdots,\omega_n)\lambda_n(\omega_1,\cdots,\omega_{n-1},d\omega_n)\cdots\lambda(d\omega_1)$$

有一有限.

若上述两积分都有限时，则 f 对 $\lambda^{(n)}$ 可积.

当积分存在时.

$$\int_{\Omega^{(n)}} f(\omega_1,\cdots,\omega_n)\,d\lambda^{(n)}$$
$$=\int_{\Omega_1}\cdots\int_{\Omega_n} f(\omega_1,\cdots,\omega_n)\lambda_n(\omega_1,\cdots,\omega_{n-1},d\omega_n)\cdots\lambda(d\omega_1)$$

注 这条定理的内容实际相当于乘积测度情形的定理 2,5 (与 $(n,\cdots,1)=(i_1,\cdots,i_k,j_1,\cdots,j_{n-k})$ 相应的一部分) 及推论 3,6 (一部分) 的总括. 其证法与上述结论的相当部分相同, 从略.

对于转移测度, 相当于推论 6 的结果是下列的

8. 定理 设 $(\Omega_i,\mathcal{F}_i)$, $i=1,2$, 为二可测空间, λ_1,λ_2 分别是 $\mathcal{F}_1$ 及 $\Omega_1\times\mathcal{F}_2$ 上的 $\sigma-$ 有限转移测度, $\lambda^{(2)}$ 为按定理 1.18 确定的测度, f 是 $\mathcal{F}_2-$ 可测函数. 若 $\int_{\Omega^{(2)}} f\,d\lambda^{(2)}$ 存在, 令

$$\mu(A):=\int_{\Omega_1}\lambda_2(\omega_1,A)\lambda_1(d\omega_1),\quad A\in\mathcal{F}_2.$$

则 μ 是 $\mathcal{F}_2$ 上的测度, 且

$$\int_{\Omega_1}\int_{\Omega_2} f(\omega_2)\lambda_2(\omega_1,d\omega_2)\lambda_1(d\omega_1)$$
$$=\int_{\Omega_2} f(\omega_2)[\int_{\Omega_1}\lambda_2(\omega_1,d\omega_2)\lambda_1(d\omega_1)]=\int_{\Omega_2} f(\omega_2)\mu(d\omega_2).$$

9. 推论 设 f 是 $\mathcal{F}_1\times\cdots\times\mathcal{F}_n$ 可测函数,

1) 若 $A_k\in\mathcal{F}_k$, $k=1,2,\cdots,n$, $fI_{A_1\times\cdots\times A_n}$ 的积分存在, 则

$$\int\limits_{A_1\times\cdots\times A_n} fd(\mu_1\times\cdots\times\mu_n)$$
$$=\int_{A_{i_n}}\left(\cdots\left(\int_{A_{i_1}} f(\omega_1,\cdots,\omega_n)\mu_{i_1}(d\omega_{i_1})\right)\cdots\right)\mu_{i_n}(d\omega_{i_n}).$$

其中 $(i_1,\cdots,i_n)$ 是 $(1,2,\cdots,n)$ 的任一置换.

2) 若 $G \in \mathcal{F}_1 \times \cdots \times \mathcal{F}_n$, fI_G 的积分存在, 则

$$\int_G f d(\mu_1 \times \cdots \times \mu_n) =$$

$$\int_{\Omega_{i_1} \times \cdots \times \Omega_{i_k}} \left(\int_{G(\omega_{i_1}, \cdots, \omega_{i_k})} f(\omega_1, \cdots, \omega_n) \right.$$

$$\left. d(\mu_{j_1} \times \cdots \times \mu_{j_{n-k}}) \right) d(\mu_{i_1} \times \cdots \times \mu_{i_k}),$$

其中 $(i_1, \cdots, i_k, j_1, \cdots, j_{n-k})$ 是 $(1, 2, \cdots, n)$ 的一个置换, $i_1 < \cdots < i_k$, $j_1 < \cdots < j_{n-k}$, 而

$$G(\omega_{i_1}, \cdots, \omega_{i_k}) = \left\{(\omega_{j_1}, \cdots, \omega_{j_{n-k}}) : (\omega_1, \cdots, \omega_n) \in G\right\}.$$

10. 推论 若 f_k 是 $(\Omega_k, \mathcal{F}_k, \mu_k)$ 上的可积函数, $k = 1, \cdots, n$, 则函数 $f(\omega_1, \cdots, \omega_n) = \prod_{k=1}^n f_k(\omega_k)$ 是对 $\mu_1 \times \cdots \times \mu_n$ 可积的函数, 且对一切 $A_k \in \mathcal{F}_k$, $k = 1, \cdots, n$,

$$\int_{A_1 \times \cdots \times A_n} f \, d(\mu_1 \times \cdots \times \mu_n) = \prod_{k=1}^n \int_{A_k} f_k \, d\mu_k.$$

Fubini 定理的应用很多, 我们在此介绍它对积分号下积分微分, L–S 积分, 重级数以及对概率论的应用.

11. 引理 设 X 是非负 r.v., F 是它的分布函数, $s \in [0, \infty)$, 则

$$\mathbf{E}X^s = s \int_0^\infty t^{s-1}[1 - F(t)] \, dt.$$

因此, 对任一 r.v. Y 有

$$\sum_{n=1}^\infty \mathbf{P}(|Y| \geq n) \leq \mathbf{E}|Y| \leq 1 + \sum_{n=1}^\infty \mathbf{P}(|Y| \geq n).$$

证明 由期望的定义及 Fubini 定理知

$$\mathbf{E}X^s = \int_{[0,\infty)} s\int_{[0,x)} t^{s-1}\,dt\mathbf{P}_X(dx)$$

$$= s\int_{[0,\infty)} t^{s-1}\Big[\int_{(t,\infty)} \mathbf{P}_X(dx)\Big]dt = s\int_0^\infty t^{s-1}[1-F(t)]dt.$$

注意到分布函数最多有可数个跳跃点，而可数集的 Lebesgue 测度为零. 所以在上述结论中 $F(t)$ 可以换成 $F(t-)$. 再令 $s=1$, $X=|Y|$ 即得

$$\mathbf{E}|Y| = \int_0^\infty \mathbf{P}(|Y|\ge t)\,dt.$$

由不等式 $\mathbf{P}(|Y|\ge n+1)\le \int_n^{n+1}\mathbf{P}(|Y|\ge t)\,dt\le \mathbf{P}(|Y|\ge n)$ 即得后一结论. □

12. 推论 设 $[a,b]$ 为一有限区间，若对一切 $\omega\in\Omega$, $f_t(\omega)$ 对 t 连续，且对任何 $(t,\omega)\in[a,b]\times\Omega$, $|f_t(\omega)|\le g(\omega)$, g 对 μ 可积，则对一切 $t\in[a,b]$

$$\int_a^t\left(\int f_s(\omega)\mu(d\omega)\right)ds = \int\left(\int_a^t f_s(\omega)\,ds\right)\mu(d\omega). \tag{1}$$

证明 若 $f_t(\omega)$ 作为 (t,ω) 的函数 $\mathcal{B}[a,b]\times\mathcal{F}$ 可测，则对任何 $t\in[a,b]$,

$$\int_{[a,t]}\left(\int_\Omega |f_s(\omega)|\,\mu(d\omega)\right)ds \le \int_{[a,t]}\left(\int_\Omega g\,d\mu\right)ds$$

$$= (t-a)\int_\Omega g\,d\mu < \infty,$$

所以由推论 3 知 (1) 成立.

今往证 $f_t(\omega)$(看成是 (t,ω) 的函数)$\mathcal{B}[a,b]\times\mathcal{F}$ 可测，令 $t_k^{(n)} = a+\dfrac{b-a}{n}k$, $k=0,\cdots,n$, 则

$$f_t^{(n)}(\omega) = \sum_{k=0}^{n-2} I_{[t_k^{(n)},t_{k+1}^{(n)})}(t)f_{t_k^{(n)}}(\omega) + I_{[t_{n-1}^{(n)},t_n^{(n)}]}(t)f_{t_k^{(n)}}(\omega)$$

为 $\mathcal{B}[a,b]\times\mathcal{F}$ 可测 (因为 $I_{[t_k^{(n)},t_{k+1}^{(n)})}(t)$, $I_{[t_{n-1}^{(n)},t_n^{(n)})}(t)$ 及 $f_{t_k^{(n)}}(\omega)$ 都是 $\mathcal{B}[a,b]\times\mathcal{F}$ 可测的). 而且

$$\lim_{n\to\infty} f_t^{(n)}(\omega)=f_t(\omega).$$

所以 $f_t(\omega)$ 是 $\mathcal{B}[a,b]\times\mathcal{F}$ 可测的. □

习题

1. 应用 Fubini 定理证明：若 $n\in\mathbb{N}$, $\mathbf{E}\xi^n$ 存在，则

$$\mathbf{E}\xi^n=n\int_0^\infty t^n[1-F(t)]\,dt-n\int_{-\infty}^0 t^{n-1}F(t)\,dt.$$

2. 设 c 为固定常数， $c>0$. 则 $\mathbf{E}|X|<\infty$ 的充要条件是

$$\sum_{n=1}^\infty \mathbf{P}(|X|\ge cn)<\infty.$$

特别是，如果对于 c 的某个值上面的级数收敛，则它对 c 的所有值也都收敛.

3. 对于任何 $r>0$, $\mathbf{E}|X|^r<\infty$ 的充要条件是

$$\sum_{n=1}^\infty n^{r-1}\mathbf{P}(|X|\ge n)<\infty.$$

4. (分部积分公式) 设 g_i, $i=1,2$, 为 $\mathcal{B}[a,b]$ 上的可测函数，而 F_i 为 $[a,b]$ 上的分布函数， $G_i(x)=\int_a^x g_i(u)\,dF_i(u)$, $x\in[a,b]$, 且

$$\int_a^b |g_i(x)|\,dF_i(x)<\infty,\qquad i=1,2.$$

则

$$\int_a^b G_1(x)g_2(x)\,dF_2(x)=G_1(b)G_2(b)-\int_a^b g_1(x)G_2(x-)\,dF_1(x).$$

其中 $\int_a^x$ 理解为 $\int_{(a,x]}$.

5. 试证定理 9.

6. 试证定理 10.

§3. 无穷维乘积概率

本节要构造任意多个概率空间的乘积概率空间，即任意多个独立 r.v. 的概率空间.

1. 乘积集的表示. 为了本节以及今后的应用，我们要给出乘积集及其元素的一种新表示法.

先看原来的表示法： $A_1\times\cdots\times A_n=\big\{(a_1,\cdots,a_n):\ a_k\in A_k,\ k=1,2,\cdots,n\big\}$. 这个表示法在定理 2.5 的叙述中已经不方便了，需要申明 $j_1,\cdots,j_{n-k}$ 的顺序. 更不方便的是无穷个集的乘积集的表示，当给定可数个集 A_n, $n\in\mathbb{N}$, 时，它们的乘积集自然表示成

$$A_1\times\cdots\times A_n\times\cdots:=\Big\{(a_1,\cdots,a_n,\cdots):\ a_n\in A_n,\ n\in\mathbb{N}\Big\}.$$

但定义不可数无穷多个集 A_t, $t\in[a,b]$, 的乘积集，就有问题了，因为 $t\in[a,b]$ 不能排序.

出路何在？

首先注意，其实 $(a_1,\cdots,a_n)$, $(a_1,\cdots,a_n,\cdots)$ 是用不着排序的，关键是每个坐标分量取自什么集？这一点在 a_k 的脚码 "k" 已经反映出来： a_k 取自 A_k. 因此 $(a_1,\cdots,a_n)$, $(a_1,\cdots,a_n,\cdots)$ 分别用集合符号 $\{a_k:\ k=1,2,\cdots,n\}$, $\{a_n:\ n\in\mathbb{N}\}$ 表示亦可，不过需要特别注明： $a_k\in A_k$. 于是，我们给出下列任意多个集的乘积集的定义及表示法.

设 I 是任意 (指标) 集，则给定集类 $\{A_i:\ i\in I\}$ 的乘积集定义为

$$(1)\qquad \prod_{i\in I}A_i:=(\text{或}\times_{k\in I}A_k:=)\Big\{\{a_i:i\in I\}:a_i\in A_i,i\in I\Big\}.$$

这种表示法不再强调乘积集的各因子集的顺序，但显然这个定义包括上面所说的情形．事实上，$A_1\times\cdots\times A_n$ 现在就写成

$$\prod_{i=1}^n A_i=\Big\{\{a_k:k=1,2,\cdots,n\}:a_k\in A_k,k=1,2,\cdots,n\Big\}.$$

这种表示包括了很广的内容，例如 I 到 X 的映射 f 也可以表成

$$\Big\{f(i):\ i\in I\Big\}.$$

但是即令是最简单的情况，也会有麻烦，例如 $\left(\dfrac{1}{2},\dfrac{1}{3}\right)\in[0,1]\times[0,1]$ 的意思是非常明确的，而 $\left\{\dfrac{1}{2},\dfrac{1}{3}\right\}\in[0,1]\times[0,1]$ 的意义却不清楚了，对于这种情形就必须写成 $\{x_1,x_2\}$，$x_1=\dfrac{1}{2}$，$x_2=\dfrac{1}{3}$，所以 (1) 的作用主要是理论讨论时方便．

为了讨论任意多个概率空间和乘积空间，还需引进一些符号，首先将乘积 $\sigma-$代数$\mathcal{F}_1\times\cdots\times\mathcal{F}_n$ 记成 $\prod_{i=1}^n\mathcal{F}_i$，将 $\mathbf{P}_1\times\cdots\times\mathbf{P}_n$ 记成 $\prod_{k=1}^n\mathbf{P}_k$．

设 I 是任意给定的无限指标集，最常用的是 $\mathbb{N}$，$[a,b]$、(a,b)，$[0,\infty)$，$[0,\infty]$、$\mathbb{R}$，$\overline{\mathbb{R}}$，$\mathbb{Z}^d$，$\mathbb{R}^d$ 等．对一切 $t\in I$，有一概率空间 $(\Omega_t,\mathcal{F}_t,\mathbf{P}_t)$．定义 Ω_t，$t\in I$，的乘积空间

$$(2)\qquad \begin{aligned}\Omega_I&:=\prod_{t\in I}\Omega_t:=\times_{t\in I}\Omega_t\\&:=\Big\{\omega_I:\ \omega_I:=\{\omega(t):\ t\in I\},\ \omega(t)\in\Omega_t,\ t\in I\Big\}\end{aligned}$$

为构造 $\mathcal{F}_t$，$t\in I$，的乘积 $\sigma-$ 代数，要借助于有限维乘积 $\sigma-$代数的概念，设 $I_N\subset I$ 为有限集，将形如

$$B_{I_N}\times\Omega_{I\setminus I_N},\qquad B_{I_N}\in\prod_{t\in I_N}\mathcal{F}_t$$

的集称为以 B_{I_N} 为底的 Ω_I 中的可测柱集．

$$(3)\qquad \mathcal{C}_I:=\Big\{B_{I_N}\times\Omega_{I\setminus I_N}:\ I_N\subset I\text{为有限集},B_{I_N}\in\prod_{t\in I_N}\mathcal{F}_t\Big\}$$

称为 Ω_I 中的可测柱集类. 称 $\sigma(\mathcal{C}_I)$ 为 $\mathcal{F}_t, t\in I$ 的乘积 $\sigma-$ 代数, 记作

$$(4)\qquad \mathcal{F}_I := \prod_{t\in I}\mathcal{F}_t := \times_{t\in I}\mathcal{F}_t := \sigma(\mathcal{C}_I).$$

我们先来证明可数无穷维乘积概率存在定理, 即

2. 定理 设 $(\Omega_n, \mathcal{F}_n, \mathbf{P}_n)$, $n\in\mathbb{N}$, 是给定的可数个概率空间, 则在 $(\Omega,\mathcal{F}) := (\Omega_{\mathbb{N}}, \mathcal{F}_{\mathbb{N}})$ 上有唯一的概率 $\mathbf{P}$ 满足条件:

$$(1)\qquad \mathbf{P}\Big(\prod_{m\in I_f}A_m\times\Omega_{\mathbb{N}\setminus I_f}\Big) = \prod_{m\in I_f}\mathbf{P}_m(A_m),$$

$$\forall I_f\subset\mathbb{N}, \text{有限}. \quad A_m\in\mathcal{F}_m,\ m\in I_f.$$

称满足 (1) 的概率 $\mathbf{P}$ 为 $\mathbf{P}_n$, $n\in\mathbb{N}$ 的乘积概率, 记作 $\mathbf{P}_{\mathbb{N}}$, 或 $\prod_{n\in\mathbb{N}}\mathbf{P}_n$, 或 $\times_{n\in\mathbb{N}}\mathbf{P}_n$; 而 $(\Omega_{\mathbb{N}}, \mathcal{F}_{\mathbb{N}}, \mathbf{P}_{\mathbb{N}})$ 称为 $(\Omega_n, \mathcal{F}_n, \mathbf{P}_n)$, $n\in\mathbb{N}$, 的乘积概率空间.

证明 易见 $\mathcal{C}_{\mathbb{N}}$ 是 $\Omega_{\mathbb{N}}$ 的一个集代数, 定理的证明想法是先在 $\mathcal{C}_{\mathbb{N}}$ 上定义一个满足 (1) 的非负值有限可加集函数 $\mathbf{P}$, 且 $\mathbf{P}(\Omega)=1$, 其次证明 $\mathbf{P}$ 在 $\emptyset$ 处连续 (这是关键), 因而 $\mathbf{P}$ 是概率, 最后应用测度扩张定理证明存在唯一性.

(一) 对一切 $C\in\mathcal{C}_{\mathbb{N}}$, 存在 I_f, 及 $A\in\prod_{k\in I_f}\mathcal{F}_k$ 使 $C=A\times\Omega_{\mathbb{N}\setminus I_f}$. 令

$$(2)\qquad \mathbf{P}(C) := \Big(\prod_{k\in I_f}\mathbf{P}_k\Big)(A).$$

今往证由 (2) 定义的 $\mathbf{P}$ 在 $\mathcal{C}_{\mathbb{N}}$ 上非负, 有限可加且 $\mathbf{P}(\Omega)=1$.

1) $\mathbf{P}$ 是 $\mathcal{C}_{\mathbb{N}}$ 上的集函数, 即 $\mathbf{P}(C)$ 与 C 的表法无关, 若

$$C = A_1\times\Omega_{\mathbb{N}\setminus I_1} = A_2\times\Omega_{\mathbb{N}\setminus I_2}, \quad A_i\in\mathcal{F}_{I_i},\ i=1,2.$$

令

$$A = A_1\times\Omega_{I_2\setminus I_1},$$

则 $A \in \mathcal{C}_{I_1 \cup I_2}$ 且

$$A = A_2 \times \Omega_{I_1 \setminus I_2}.$$

于是由 $\mathbf{P}_n(\Omega_n) = 1$ 知

$$\Big(\prod_{k\in I_1} \mathbf{P}_k\Big)(A_1) = \Big(\prod_{k\in I_1\cup I_2} \mathbf{P}_k\Big)(A) = \Big(\prod_{k\in I_2} \mathbf{P}_k\Big)(A_2),$$

即 (2) 的值与 C 的表法无关.

2) 显然 $\mathbf{P}$ 在 $\mathcal{C}_{\mathbb{N}}$ 上非负且 $\mathbf{P}(\Omega_{\mathbb{N}}) = 1$, $\mathbf{P}$ 满足 (1).

3) $\mathbf{P}$ 在 $\mathcal{C}_{\mathbb{N}}$ 上有限可加, 由于 $\mathcal{C}_{\mathbb{N}}$ 是集代数, 所以只需证: $\mathbf{P}$ 在 $\mathcal{C}_{\mathbb{N}}$ 上可加, 事实上, 若 $A_1, A_2 \in \mathcal{C}_{\mathbb{N}}$, 且 $A_1 \cap A_2 = \varnothing$, 则存在 $I_1 \subset \mathbb{N}$ 为有限集, $A'_k \in \mathcal{F}_{I_k}$, $k = 1, 2$, 使

$$A_k = A'_k \times \Omega_{\mathbb{N}\setminus I_1}.$$

由 $A_1 \cap A_2 = \varnothing$ 知 $A'_1 \cap A'_2 = \varnothing$, 故由 (2)

$$\begin{aligned}\mathbf{P}(A_1 \cup A_2) &= \mathbf{P}_{I_1}(A'_1 \cup A'_2) = \mathbf{P}_{I_1}(A'_1) + \mathbf{P}_{I_1}(A'_2)\\ &= \mathbf{P}(A_1) + \mathbf{P}(A_2).\end{aligned}$$

(二) 证明 $\mathbf{P}$ 在 $\varnothing$ 处连续, 即证明 "若 $\{A_n\} \subset \mathcal{C}_{\mathbb{N}}$, $A_n \downarrow \varnothing$, 则 $\mathbf{P}(A_n) \to 0$". 为此, 我们证明它的等价命题:

"若 $\{A_n\} \subset \mathcal{C}_{\mathbb{N}}$, $A_n \downarrow$, 且存在 $\varepsilon > 0$, 使对一切 $n \in \mathbb{N}$, $\mathbf{P}(A_n) \geq \varepsilon$, 则 $\bigcap_{n=1}^{\infty} A_n \neq \varnothing$."

我们的办法是证明存在一个点 $(\bar\omega_1, \bar\omega_2, \cdots)$, 属于一切 A_n, (想法是考虑 $B_1^{(n)} := \{\omega_1 \in \Omega_1 : A_n(\omega_1)$ 的 "概率较大"$\}$ 而往证 $\bigcap_{n\in\mathbb{N}} B_1^{(n)} \neq \varnothing, \cdots$) 为此令

$$\Omega_{(>n)} := \prod_{k=n+1}^{\infty} \mathcal{F}_k,$$

$$\mathcal{C}_{(>n)} := \Big\{B \times \Omega_{(>n+m)}:\ B \in \prod_{k=n+1}^{n+m} \mathcal{F}_k,\ m \in \mathbb{N}\Big\},$$

$$\mathcal{F}_{(>n)} := \prod_{k=n+1}^{\infty} \mathcal{F}_k.$$

仿 (一) 定义集函数

$$\mathbf{P}_{(>n)}(B \times \Omega_{(>n+m)}) := \Big(\prod_{k=n+1}^{n+m} \mathbf{P}_k\Big)(B), \quad \forall\, B \in \prod_{k=n+1}^{n+m} \mathcal{F}_k,$$

则和 (一) 的证明一样，可证 $\mathbf{P}_{(>n)}$ 是 $\mathcal{C}_{(>n)}$ 上非负有限可加集函数，且

$$\mathbf{P}_{(>n)}\Big(\prod_{k=n+1}^{n+m} A_k \times \Omega_{(>n+m)}\Big) = \prod_{k=n+1}^{n+m} \mathbf{P}_k(A_k).$$

此外还可证明

$$\mathbf{P}(A) = \int_{\prod_{k=1}^n \Omega_k} \mathbf{P}_{(>n)}(A(\omega_1, \cdots, \omega_n))\, d\Big(\prod_{k=1}^n \mathbf{P}_k\Big), \quad \forall\, n \in \mathbb{N}, \quad A \in \mathcal{C}_{\mathbb{N}}, \tag{3}$$

$$\begin{aligned}&\mathbf{P}_{(>n)}(A)\\ &= \int_{\prod_{k=n+1}^{n+m} \Omega_k} \mathbf{P}_{(>n+m)}(A(\omega_{n+1}, \cdots, \omega_{n+m}))\, d\Big(\prod_{k=n+1}^{n+m} \mathbf{P}_k\Big),\\ &\qquad \forall\, n, m \in \mathbb{N} \qquad \forall\, A \in \mathcal{C}_{(>n)}.\end{aligned} \tag{4}$$

事实上，对一切 $n \in \mathbb{N}$、$A \in \mathcal{C}_{\mathbb{N}}$，可设 $A = A' \times \Omega_{(>m)}$，$m > n$，$A' \in \prod_{k=1}^m \mathcal{F}_k$，于是

$$\begin{aligned}&A(\omega_1, \cdots, \omega_n) = A'(\omega_1, \cdots, \omega_n) \times \Omega_{(>m)},\\ &A'(\omega_1, \cdots, \omega_n) \in \prod_{k=n+1}^m \mathcal{F}_k,\\ &\mathbf{P}(A) = \Big(\prod_{k=1}^m \mathbf{P}_k\Big)(A') = \Big(\prod_{k=1}^n \mathbf{P}_k\Big) \times \Big(\prod_{k=n+1}^m \mathbf{P}_k\Big)(A')\\ &= \int_{\prod_{k=1}^n \Omega_k} \Big(\prod_{k=n+1}^m \mathbf{P}_k\Big)(A'(\omega_1, \cdots, \omega_n))\, d\Big(\prod_{k=1}^n \mathbf{P}_k\Big)\\ &= \int_{\prod_{k=1}^n \Omega_k} \mathbf{P}_{(>n)}(A(\omega_1, \cdots, \omega_n))\, d\Big(\prod_{k=1}^n \mathbf{P}_k\Big)\end{aligned}$$

故 (3) 获证，同样可证 (4).

现在我们可以实现前面的想法：令

$$B_1^{(n)} := \Big\{\omega_1 \in \Omega_1 :\ \mathbf{P}_{(>1)}(A_n(\omega_1)) \geq \frac{\varepsilon}{2}\Big\}.$$

显然 $A_n(\omega_1) \supset A_{n+1}(\omega_1)$，因而

$$\mathbf{P}_{(>1)}(A_n(\omega_1)) \geq \mathbf{P}_{(>1)}(A_{n+1}(\omega_1)),$$

故 $B_1^{(n)}\downarrow$. 此外由 (3) 的证明知 $\mathbf{P}_{(>1)}(A_n(\omega_1))$ 是 $\mathcal{F}_1$- 可测函数. 因此 $B_1^{(n)}\in\mathcal{F}_1$, 对一切 $n\in\mathbb{N}$. 而对一切 $n\in\mathbb{N}$

$$\begin{aligned}\varepsilon\leq\mathbf{P}(A_n)&=\int_{\Omega_1}\mathbf{P}_{(>1)}(A_n(\omega_1))\,d\mathbf{P}_1\\&=\int_{B_1^{(n)}}\mathbf{P}_{(>1)}(A_n(\omega_1))d\mathbf{P}_1+\int_{(B_1^{(n)})^c}\mathbf{P}_{(>1)}(A_n(\omega_1))d\mathbf{P}_1\\&\leq\mathbf{P}_1(B_1^{(n)})+\frac{\varepsilon}{2}.\end{aligned}$$

即得

$$\mathbf{P}_1(B_1^{(n)})\geq\frac{\varepsilon}{2},\qquad\forall\, n\in\mathbb{N}. \tag{5}$$

由 $\mathbf{P}_1$ 是 $\mathcal{F}_1$ 上的概率知它在 $\emptyset$ 处连续, 故由 (5) 知 $\bigcap_{n\in\mathbb{N}}B_1^{(n)}\neq\emptyset$. 即存在 $\bar{\omega}_1\in\bigcap_{n\in\mathbb{N}}B_1^{(n)}$, 亦即存在 $\bar{\omega}_1\in\Omega_1$, 使 $\mathbf{P}_{(>1)}(A_n(\bar{\omega}_1))\geq\varepsilon/2$. 其次, $A_n(\bar{\omega}_1)\in\mathcal{C}_{(>1)}$, $\mathbf{P}_{(>1)}(A_n(\bar{\omega}_1))\geq\varepsilon/2$, $n\in\mathbb{N}$. 仿前, 令

$$B_2^{(n)}:=\Big\{\omega_2\in\Omega_2:\ \mathbf{P}_{(>2)}(A_n(\bar{\omega}_1,\omega_2))\geq\frac{\varepsilon}{2^2}\Big\}.$$

注意 $A_n(\bar{\omega}_1,\omega_2)=(A_n(\bar{\omega}_1))(\omega_2)$, 完全重复前面的论证 (在应用 (3) 处换成应用 (4)) 知存在 $\bar{\omega}_2\in\Omega_2$, 使对一切 $n\in\mathbb{N}$,

$$\mathbf{P}_{(>n)}(A_n(\bar{\omega}_1,\bar{\omega}_2))\geq\varepsilon/2^2.$$

用同样的方法及归纳法可证

$$\exists\,(\bar{\omega}_1,\bar{\omega}_2,\cdots)\in\Omega,\ \forall\, n\in\mathbb{N},\quad m\in\mathbb{N},$$
$$\text{有 }\mathbf{P}_{(>n)}(A_m(\bar{\omega}_1,\cdots,\bar{\omega}_n))\geq\frac{\varepsilon}{2^n}.$$

往证 $(\bar{\omega}_1,\bar{\omega}_2,\cdots,)\in\bigcap_{m=1}^{\infty}A_m$.

对一切 $m\in\mathbb{N}$, 由 $A_m\in\mathcal{C}_{\infty}$ 知存在 $n(m)\in\mathbb{N}$, $A'_m\in\prod_{k=1}^{n(m)}\mathcal{F}_k$ 使

$$A_m=A'_m\times\Omega_{(>n(m))}.$$

而

$$\mathbf{P}_{(>n(m))}(A_m(\bar{\omega}_1,\cdots,\bar{\omega}_{n(m)})) \geq \frac{\varepsilon}{2^{n(m)}} > 0.$$

因此,

$$\begin{aligned}\emptyset \neq\ & A_m(\bar{\omega}_1,\cdots,\bar{\omega}_{n(m)}) \\ &= \begin{cases} \Omega_{(>n(m))}, & 当(\bar{\omega}_1,\cdots,\bar{\omega}_{n(m)}) \in A'_m, \\ \emptyset, & 当(\bar{\omega}_1,\cdots,\bar{\omega}_{n(m)}) \notin A'_m. \end{cases}\end{aligned}$$

故 $(\bar{\omega}_1,\cdots,\bar{\omega}_{n(m)}) \in A'_m$. 因而对一切 $m \in \mathbb{N}$, 对一切 $(\omega_{n(m)+1},\omega_{n(m)+2},\cdots)$ 有

$$(\bar{\omega}_1,\cdots,\bar{\omega}_{n(m)},\omega_{n(m)+1},\omega_{n(m)+2},\cdots) \in A_m$$

特别 $(\bar{\omega}_1,\bar{\omega}_2,\cdots) \in \bigcap_{n=1}^{\infty} A_n$.

这样就证明 $\mathbf{P}$ 在 $\emptyset$ 连续. 因而 $\mathbf{P}$ 是 $\mathcal{C}_{\mathbb{N}}$ 上的概率且满足 (2).

(三) 由于 $\mathcal{C}_{\mathbb{N}}$ 是集代数, 因而自然是半集代数, 而 $\sigma(\mathcal{C}_{\mathbb{N}}) = \mathcal{F}_{\mathbb{N}}$. 故由测度扩张定理知满足 (2) (即在 $\mathcal{C}_{\mathbb{N}}$ 上相同) 的概率在 $\mathcal{F}_{\mathbb{N}}$ 上的扩张是存在且唯一的.

故由 (一) 至 (三) 知定理获证. □

现在介绍 Tulcea 定理. 它是可数维无穷乘积概率定理的推广, 即由转移概率构造无穷乘积空间上的概率的定理, 其证明方法也几乎完全相同.

3. 定理 (Tulcea 定理) 设 $(\Omega_n, \mathcal{F}_n)$, $n \in \mathbb{N}$, 是可测空间,

$$\Omega^{(n)} := \prod_{k=1}^{n} \Omega_k, \qquad \mathcal{F}^{(n)} := \prod_{k=1}^{n} \mathcal{F}_k;$$

$$\Omega_{\mathbb{N}} := \prod_{n\in\mathbb{N}} \Omega_n, \qquad \mathcal{F}_{\mathbb{N}} := \prod_{n\in\mathbb{N}} \mathcal{F}_n.$$

$\mathbf{P}_n(\omega^{(n-1)}, A)$, $\omega^{(n-1)} \in \Omega^{(n-1)}$, $A \in \mathcal{F}_n$, 是 $\Omega^{(n-1)} \times \mathcal{F}_n$, $n \geq 2$, 上的转移概率, $\mathbf{P}_1$ 是 $\mathcal{F}_1$ 上的概率. 则在 $\mathcal{F}_{\mathbb{N}}$ 上有唯一的概率 $\mathbf{P}_{\mathbb{N}}$ 存在, 使

$$\mathbf{P}_{\mathbb{N}}(C(B)) := \mathbf{P}^{(n)}(B), \qquad B \in \mathcal{F}^{(n)},$$

其中 $C(B)$ 表示 $\mathcal{F}_{\mathbb{N}}$ 中以 $B \in \mathcal{F}^{(n)}$ 为底的柱集， $\mathbf{P}^{(n)}$ 即为以 $\mathbf{P}_k$ 代换 λ_k 而由定理 1.18 确定的 $\lambda^{(n)}$.

证明 首先注意由一切 $C(B)$, $B \in \mathcal{F}^{(n)}$, $n \in \mathbb{N}$, 组成的集类是集代数 $\mathcal{C}_{\mathbb{N}}$. 和定理 2 的证明一样， $\mathbf{P}_{\mathbb{N}}$ 在 $\mathcal{C}_{\mathbb{N}}$ 上定义了一个有限可加概率. 然后完全类似于定理 2 的证明 (二) 可以证明 $\mathbf{P}_{\mathbb{N}}$ 在 $\emptyset$ 处连续 (此事留给读者作为练习)，因而 $\mathbf{P}_{\mathbb{N}}$ 是概率. 由测度扩张定理即得本定理. □

在可数无穷维乘积概率的构造问题解决之后，不可数无穷维的乘积概率构造可以很容易地归结为可数维的情形，这是由于下列的

4. 引理 设 I 为任一不可数集. $(\Omega_t, \mathcal{F}_t), t \in I$, 是一族可测空间，则 Ω_I 中 $\mathcal{F}_t, t \in I$, 的乘积 $\sigma-$ 代数

$$\begin{aligned}\mathcal{F}_I &= \bigcup_{\substack{I_c \subset I \\ I_c \text{ 可数}}} \left\{A_{I_c} \times \Omega_{I \setminus I_c} : A_{I_c} \in \mathcal{F}_{I_c}\right\} \\ &= \bigcup_{\substack{I_c \subset I \\ I_c \text{ 可数}}} \mathcal{F}_{I_c} \times \Omega_{I \setminus I_c}.\end{aligned}$$

证明留给读者 (注意 "可数个可数集的并仍然可数").

5. 定理 (无穷乘积概率存在定理) 设 $(\Omega_t, \mathcal{F}_t, \mathbf{P}_t), t \in I$, 是一族概率空间，则在它们的乘积可测空间 $(\Omega_I, \mathcal{F}_I)$ 上存在一个唯一的概率 $\mathbf{P}_I$ 使对一切 $I_f \subset I$ 为有限集，对任何 $A_t \in \mathcal{F}_t, t \in I_f$, 有

$$\mathbf{P}_I\Big(\prod\nolimits_{t \in I_f} A_t \times \Omega_{I \setminus I_f}\Big) = \prod\nolimits_{t \in I_f} \mathbf{P}_t(A_t). \tag{1}$$

$\mathbf{P}_I$ 称为 $\mathbf{P}_t, t \in I$, 的乘积概率测度，有时记作 $\prod\limits_{t \in I} \mathbf{P}_t$ 或 $\underset{t \in I}{\times} \mathbf{P}_t$. $(\Omega_I, \mathcal{F}_I, \mathbf{P}_I)$ 称为 $(\Omega_t, \mathcal{F}_t, \mathbf{P}_t), t \in I$, 的乘积概率空间.

证明 当 I 为有限或可数集时，前面已证，只需考虑 I 为不可数的情形.

对一切 $A \in \mathcal{F}_I$, 由引理 4 知存在 $I_c \subset I$ 可数集，使

$$A = A_{I_c} \times \Omega_{I \setminus I_c}, \qquad A_{I_c} \in \mathcal{F}_{I_c},$$

由定理 2 知在 $(\Omega_{I_c}, \mathcal{F}_{I_c})$ 上存在唯一的无穷乘积概率测度 $\mathbf{P}_{I_c}$. 定义

$$\mathbf{P}_I(A) = \mathbf{P}_{I_c}(A_{I_c}). \tag{2}$$

由此不难证明：$\mathbf{P}_I(A)$ 的值与 A 的表法无关，$\mathbf{P}_I$ 具有 σ–可加性，由测度扩张定理知是唯一的. □

象有限维乘积测度定理对有限个 a.e. 的独立的应用一样，无穷乘积概率定理对于证明无穷个独立 a.e. 的存在以及无穷个独立 r.v. 的性质是很有用的，为了叙述这些应用，先给出

6. 定义 设 I 为任一指标集，$(\Omega, \mathcal{F}, \mathbf{P})$ 为一概率空间，$\{A_i : i \in I\}$, 若对一切 $n \in \mathbb{N}$, $\{i_1, \cdots, i_n\} \subset I$ 有

$$\mathbf{P}\Big(\bigcap_{k=1}^{n} A_{i_k}\Big) = \prod_{k=1}^{n} \mathbf{P}(A_{i_k}). \tag{1}$$

则称事件 $A_i, i \in I$, 独立.

若对一切 $i \in I$, $\mathcal{C}_i \subset \Omega$, 满足条件：对一切 $A_i \in \mathcal{C}_i$, $i \in I$ 都有事件 $A_i, i \in I$, 独立，则称事件类 $\mathcal{C}_i, i \in I$, 独立.

设对一切 $i \in I$, X_i 是 $m_i(\in \mathbb{N})$ 维实 r.v.，若事件类

$$\mathcal{C}_i := \big\{\{X_i \le x\} :\ x \in \mathbb{R}^{m_i}\big\}, \qquad i \in I,$$

独立，则称 r.v. $X_i, i \in I$, 独立.

和推论 1.19、1.20 一样有下列的

7. 推论 设 X_i 是 $(\Omega_i, \mathcal{F}_i, \mathbf{P}_i)$ 上的 m_i 维实 r.v.，$i \in I$, 则在乘积概率空间 $(\Omega_I, \mathcal{F}_I, \mathbf{P}_I)$ 上定义 r.v.

$$Y_i : Y_i(\omega_I) = X_i(\omega_i),\ \forall \omega_I = \{\omega_i : i \in I\} \in \Omega_I,\ i \in I,$$

则 Y_i, $i \in I$, 是独立 r.v.，且 Y_i 与 X_i 同分布， $i \in I$.

证明　由推论 1.19 的证明知对一切 $n \in \mathbb{N}$, $\{i_1, \cdots, i_n\} \subset I$, $Y_{i_1}, \cdots, Y_{i_n}$ 独立，且对一切 $i \in I$, Y_i 与 X_i 同分布，故 Y_i, $i \in I$, 独立. □

应用乘积空间的概念及定理 5 (或单调类定理) 还可以证明下列

8. 定理　设 I 是任一指标集，$(\Omega, \mathcal{F}, \mathbf{P})$ 是一个概率空间.

1) 若 $\pi-$ 系 $\mathcal{C}_i \subset \mathcal{F}$, $i \in I$, 独立，则子 $\sigma-$ 代数 $\sigma(\mathcal{C}_i)$, $i \in I$, 独立.

2) 若 $\mathcal{F}_i$, $i \in I$, 独立，$I = \bigcup_{t \in T} I_t$ 且 I_t, $t \in T$, 两两不交，则乘积 $\sigma-$ 代数 $\mathcal{F}_{I_t}$, $t \in T$, 独立.

定理 8 包括很广的内容. 例如若 X_i, $i \in I$, 独立，$I = \bigcup_{t \in T} I_t$, I_t, $t \in I$, 两两不交，f_t 是 $(\mathbb{R}_{I_t}, \mathcal{B}_{I_t})$ 到 $(\mathbb{R}, \mathcal{B})$ 的可测函数，则 $f_t(\{X_i:\ i \in I_t\})$, $t \in T$, 是独立 r.v. .

9. 注　应用推论 7 固然可以构造具有任意给定的分布族的独立 r.v. 族，但是对一些具体情形，人们也常采用一些具体构造的办法. 例如可以用以下办法构的办法. 例如如可以用以下办法构造 Bernoulli 序列 $\{X_n:\ n \in \mathbb{N}\}$ (即 X_n, $n \in \mathbb{N}$, 独立，且每一 X_n 具有两点分布 $\begin{pmatrix} 0 & 1 \\ p & 1-p \end{pmatrix}$, $p \in (0,1)$). 考虑 $([0,1), \mathcal{B}[0,1))$ 在 $[0,1)$ 上定义 X_n 如下：从图形上看，将区间 $[0,1)$ 分成

$$\delta_0 := [0, p), \qquad \delta_1 := [p, 1).$$

令 $X_1 : [0,1) \to \{0,1\}$, 且在 δ_i 上取值 i, $i = 0, 1$, 即 $X_1(\delta_0) = \{0\}$, $X_1(\delta_1) = \{1\}$. 再令

$$\delta_{00} := [0, p^2), \qquad \delta_{01} := [p^2, p),$$

$$\delta_{10} := [p, p + p(1-p)), \quad \delta_{11} := [p + p(1-p), 1),$$

即将 δ_k 按长度比 $p : (1-p)$ 分成 δ_{k0}, δ_{k1}. 再令

$$X_2(\delta_{ki}) = \{i\}, \qquad i = 0, 1, \quad k = 0, 1.$$

设 $\delta_{k_1\cdots k_n}$, $k_j\in\{0,1\}$, $j=1,\cdots,m$ 及 X_n 已经定义，则将每一 $\delta_{k_1\cdots k_n}$ 按长度比 $p:(1-p)$ 分成 $\delta_{k_1\cdots k_n 0}$、$\delta_{k_1\cdots k_n 1}$，且定义

$$X_{n+1}:[0,1)\to\{0,1\},\ X_{n+1}(\delta_{k_1\cdots k_n i})=\{i\},\ i=0,1.$$

不难证明：这样定义的 X_n, $n\in\mathbb{N}$, 是 $([0,1),\mathcal{B}[0,1),\lambda)$ 上的独立 r.v.，且每一 X_n 的分布是 $\begin{pmatrix}0 & 1\\ p & 1-p\end{pmatrix}$.

当 $p=\frac{1}{2}$ 时，$X_n(t)$, $n\in\mathbb{N}$, 即为 $t\in[0,1)$ 的二进位展开的小数点后第 n 个数码，即

$$t=\sum_{n=1}^{\infty}\frac{X_n(t)}{2^n},\qquad t\in[0,1).$$

上面的 $X_n(t)$ 也可如下分析地定义：令变换

$$T_p(t):=\begin{cases} t/p, & 0\le t<p,\\ (t-p)/(1-p), & p\le t<1,\end{cases}\qquad t\in[0,1)$$

$$T_p^0(t)=t,\quad T_p^n(t)=T_p(T_p^{n-1}(t)),\qquad n\in\mathbb{N},$$

$$X_n(t):=I_{[p,1)}(T_p^{n-1}(t)),\quad t\in[0,1).$$

$X_n(t)$, $n\in\mathbb{N}$, 的两种定义的等价性及其独立性留给读者作为习题.

10. 注 由于推论 7，概率论中对独立变量序列 (包括 **独立试验序列**) 的各种讨论才有了理论 (逻辑) 基础，下列我们再给出 **Possion 过程** 及其他例子.

例1 设 N 是取正整数值的 r.v.，$X_n, n\in\mathbb{N}$, 是同分布 r.v.，$\mathbf{E}X_n<\infty$, 且 N, X_n, $n\in\mathbb{N}$, 独立,

$$Y:=\sum_{n=1}^{N}X_n$$

则

$$\mathbf{E}Y=\mathbf{E}X_1\cdot\mathbf{E}N.\qquad (\text{Wald 等式})$$

证明　设 $X=(X_1,\cdots)$, 则 X 与 N 独立，因而 $\mathbf{P}_{(X,N)}=\mathbf{P}_X\times\mathbf{P}_N$, 其中 $\mathbf{P}_X$ 表示 X 的分布，即 $\mathbf{P}_X(B)=\mathbf{P}(X\in B)$, 对一切 $B\in\mathcal{B}_{\mathbb{N}}$.

$$\begin{aligned}\mathbf{E}Y&=\int\sum_{n=1}^{N(\omega)}X_n(\omega)\,\mathbf{P}(d\omega)\\&=\int\Big(\int\sum_{n=1}^{m}x_n\,\mathbf{P}_X(dx)\Big)\mathbf{P}_N(dm)\\&=\int\mathbf{E}\Big(\sum_{n=1}^{m}X_n\Big)\,\mathbf{P}_N(dm)=\int m\mathbf{E}X_1\,\mathbf{P}_N(dm)\\&=\mathbf{E}X_1\cdot\mathbf{E}N.\end{aligned}$$

以下两例曾在第五章 §3 习题 12,13 中讨论过.

例2　设 T_n, $n\in\mathbb{N}$, i.i.d.(表示独立同分布) 且 $T_n\sim\mathbf{E}(\lambda)$ (以 λ 为参数的指数分布). 令

$$N(t):=\sup\left\{n\in\mathbb{N}:\ \textstyle\sum_{k=1}^{n}T_k\le t\right\}$$

则 $\{N(t):\ t\geqslant 0\}$ 为一具参数 λ 的 Poisson 过程，即

1) $N(0)=0$.

2) 对一切 $\{t_0,\cdots,t_m\}\subset\mathbb{R}_+$, $0=t_0<t_1<\cdots<t_m$, $N(t_k)-N(t_{k-1})$, $k=1,\cdots,m$, 独立,

3) 对一切 $t>s\ge 0$, $N(t)-N(s)\sim\mathcal{P}(\lambda(t-s))$.

证明　1) 显然，为证 2), 3), 任取 $0=t_0<t_1<\cdots<t_m<\infty$, 计算 (设 $n_\ell\ge 0$, $r_\ell:=n_1+\cdots+n_\ell$, $\ell=1,2,\cdots,m$)

$$\begin{aligned}\mathbf{P}:=&\mathbf{P}(N(t_1)-N(t_0)=n_1,\cdots,N(t_m)-N(t_{m-1})=n_m)\\=&\mathbf{P}(N(t_1)=r_1,\cdots,N(t_m)=r_m)\\=&\mathbf{P}\Big(\sum_{k=1}^{r_\ell}T_k\le t_\ell<\sum_{k=1}^{r_\ell+1}T_k,\ell=1,\cdots,m\Big)\end{aligned}\tag{4}$$

以下为易于理解分情形计算:

1) $n_1 = n_2 = \cdots = n_m = 0$, 则

$$\mathbf{P} = \mathbf{P}(t_m < T_1) = \int_{t_m < s_1} \lambda exp\{-\lambda s_1\}\, ds_1 = exp\{-\lambda t_m\}$$

$$= \prod_{\ell=1}^{m} exp\{-\lambda(t_\ell - t_{\ell-1})\}\frac{[\lambda(t_\ell - t_{\ell-1})]^{n_\ell}}{n_\ell !}.$$

2) 存在 $p \geq 2$, $n_1 = \cdots = n_{p-1} = 0$, $n_p \geq 1$, $n_{p+1} = \cdots = n_m = 0$, 则

$$\mathbf{P} = \mathbf{P}(t_{p-1} < T_p, \sum_{k=1}^{n_p} T_k \leqslant t_p,\ t_m < \sum_{k=1}^{n_p+1} T_k)$$

$$= \int \cdots \int_{\substack{t_{p-1} < s_1 \\ \sum_{k=1}^{n_p} s_k \leq t_p, \\ \sum_{k=1}^{n_p+1} s_k > t_m}} \lambda^{n_p + 1} \exp\left[-\lambda \sum_{k=1}^{n_p+1} s_k\right] ds_1 \cdots ds_{n_p+1}$$

$$= \int \cdots \int_{\sum_{k=1}^{n_p} s_k \leq t_p, t_{p-1} < s_1} \lambda^{n_p} \exp\left[-\lambda \sum_{k=1}^{n_p} s_k\right] ds_1 \cdots ds_{n_p}$$

$$\times \int_{t_m - \sum_{k=1}^{n_p} s_k}^{\infty} \lambda \exp\left[-\lambda s_{n_p+1}\right] ds_{n_p+1}$$

$$= exp\{-\lambda t_m\} \int \cdots \int_{\substack{\sum_{k=1}^{n_p} s_k \leq t_p \\ t_{p-1} < s_1}} \lambda^{n_p} ds_1 \cdots ds_{n_p}$$

$$= exp\{-\lambda t_m\} \int \cdots \int_{\sum_{k=1}^{n_p} u_k \leq t_p - t_{p-1}} \lambda^{n_p} du_1 \cdots du_{n_p}$$

$$= exp\{-\lambda t_m\} \frac{[\lambda(t_p - t_{p-1})]^{n_p}}{n_p !}$$

$$= \prod_{k=1}^{m} exp\{-\lambda(t_\ell - t_{\ell-1})\} \frac{[\lambda(t_\ell - t_{\ell-1})]^{n_\ell}}{n_\ell !}$$

3) 设 $p_0 = 0$, $i = 0, 1, \cdots, j-1$, $n_{p_i+1} = \cdots = n_{p_{i+1}-1} = 0$,

$n_{p_{i+1}} \geq 1$, $n_{p_j+1} = \cdots = n_m = 0$, 于是由 (4)

$$\mathbf{P} = \int\limits_{\substack{t_{p_{i-1}} < \sum_{k=1}^{r_{p_i-1}+1} s_k, i=1,\cdots,j \\ \sum_{k=1}^{r_{p_1}-1} s_k \leq t_{p_i}, i=1,2,\cdots,j}} \cdots \int \lambda^{r_{p_j}} \exp\left[-\lambda \sum_{k=1}^{r_{p_j}} s_k\right] ds_1 \cdots ds_{n_{p_1}+\cdots+n_{p_j}}$$

$$\int\limits_{t_m - \sum_{k=1}^{r_{p_i}} s_k}^{\infty} \lambda \exp\left[-\lambda s_{r_{p_j}} + 1\right] \cdot ds_{n_{p_1}+\cdots+n_{p_j}+1}$$

$$= exp\{-\lambda t_m\}\lambda^{r_{p_j}} \int\limits_{\substack{t_{p_{i-1}} < \sum_{k=1}^{r_{p_i}+1} s_k \\ \sum_{k=1}^{r_{p_i}} s_k \leq t_{p_i}}} \cdots \int ds_1 \cdots d_{s_{n_{p_1}+\cdots+n_{p_j}}},$$

令 $r_{p_i} = \ell_i$, $i = 1, 2, \cdots, j$, $\ell_0 = 0$, 并作变换

$$u_k = \begin{cases} \sum_{k=1}^{\ell_i+1} s_k - t_{p_i-1}, & \text{当} k = r_i + 1, \\ s_k, & \text{其它} \end{cases}$$

则

$$\mathbf{P} = exp\{-\lambda t_m\}\lambda^{\ell_j} \prod_{i=1}^{j} \int\limits_{\sum_{k=\ell_{i-1}+1}^{\ell_i} u_k \leq t_{p_i} - t_{p_i-1}} \cdots \int du_{\ell_{i-1}+1} \cdots du_{\ell_i}$$

$$= exp\{-\lambda t_m\} \prod_{i=1}^{j} \frac{[\lambda(t_{p_i} - t_{p_i-1})]^{\ell_i - \ell_{i-1}}}{\ell_i!}$$

$$= \prod_{k=1}^{m} exp\{-\lambda(t_k - t_{k-1})\} \frac{[\lambda(t_k - t_{k-1})]^{n_k}}{n_k!}$$

由此即知 (2), (3) 同时获证.

例3 . 设 $X_1, X_2, \cdots, Y_1, Y_2, \cdots$, 是独立 r.v. , $X_k \sim \mathbf{E}(\lambda)$,

Y 遵从集中在 $(0,\infty]$ 上的分布 μ, 令 $N(t)$ 为满足条件

$$\{N(t)=n\}=\Big\{\sum_{k=1}^{n}(X_k+Y_k)\le t,\sum_{k=1}^{n}(X_k+Y_k)+X_{n+1}>t\Big\}.$$

的 r.v. , 则

$$\mathbf{P}(N(t)=n)=\int_{(0,1]}\frac{[\lambda(t-x)]^n}{n!}exp\{-\lambda(t-x)\}\mu^{n*}(dx)$$

其中 μ^{n*} 表示 μ 的 n 重卷积, 即

$$\mu^{n*}((0,t]):=$$
$$(\mu\times\cdots\times\mu)\big(\{(x_1,\cdots,x_n)\in\mathbb{R}_+^n:\ x_1+\cdots+x_n\le t\}\big)$$

由此

$$\begin{aligned}p(t)&=\sum_{n=0}^{\infty}\mathbf{P}(N(t)=n)\\&=exp\{-\lambda t\}+\sum_{n=1}^{\infty}\int_{(0,t]}\frac{[\lambda(t-x)]^n}{n!}exp\{-\lambda(t-x)\}\,\mu^{n*}(dx).\end{aligned}$$

(**注** 函数 $p(t)$ 在再生 (regenerative) 现象中称为 p–函数)

证明

$$\begin{aligned}&\mathbf{P}(N(t)=n)\\&=\mathbf{P}\left(\sum_{k=1}^{n}(X_k+Y_k)\le t,\ \sum_{k=1}^{n}(X_k+Y_k)+X_{n+1}>t\right)\\&=\underset{\substack{\sum_{k=1}^{n}(x_k+y_k)\le t\\ \sum_{k=1}^{n}(x_k+y_k)+x_{n+1}>t\\ x_k\ge0,y_k\ge0}}{\int\cdots\int}\lambda^{n-1}exp\{-\lambda\sum_{k=1}^{n}x_k\}dx_1\cdots dx_{n+1}\\&\qquad\mu(dy_1)\cdots\mu(dy_n)\end{aligned}$$

记 μ_Y 为 $Y=(Y_1,\cdots,Y_n)$ 的分布, $|y|=\sum_{k=1}^{n}y_k$, 则上式

$$= \int_{|y|\le t} \mu_Y(dy) \int_{\sum_{k=1}^n x_k \le t-|y|} \lambda^n exp\{-\lambda \sum_{k=1}^n x_k\} dx_1 \cdots dx_n$$

$$\times \int_{x_{n+1} > t - \sum_1^n x_k - |y|} \lambda exp\{-\lambda x_{n+1}\} dx_{n+1}$$

$$= \int_{|y|\le t} \mu_Y(dy) \int_{\sum_{k=1}^n x_k \le t-|y|} \lambda^n exp\{-\lambda (t-|y|)\} dx_1 \cdots dx_n$$

$$= \int_{|y|\le t} \frac{[\lambda(t-|y|)]^n}{n!} exp\{-\lambda(t-|y|)\} \mu_Y(dy)$$

$$= \int_{(0,t]} \frac{[\lambda(t-x)]^n}{n!} exp\{-\lambda(t-x)\} \mu^{n*}(dx).$$

习题

证明定理 3 中定义的 $\mathbf{P}_{\mathbb{N}}$ 在 $\emptyset$ 处连续.

第七章 不定积分与条件期望

符号测度的分解以及不定积分的刻划是测度论中重要论题，不定积分的 Radon Nikodym 定理则是概率论中一般条件期望的数学基础. 而条件期望是现代概率论的最重要的基本概念之一. 本章的目的就是介绍这两方面的内容.

§1. 符号测度的分解

本节的目的主要是证明符号测度的分解定理 (定理 5), 然后介绍经典的有限变差函数的一些性质.

1. 定义 设 $\mathcal{A}$ 是 Ω 中的一个集代数， $\varphi: \mathcal{A} \to \overline{\mathbb{R}}$ 称为 $\mathcal{A}$ 上的 **有限可加集函数**, 如果 $\varphi(\emptyset) = 0$ 且

$$\varphi(\bigcup_{k=1}^{n} A_k) = \sum_{k=1}^{n} \varphi(A_k)$$

对一切两两不交的 $A_k \in \mathcal{A}$, $k = 1, \cdots, n$, 成立. 称 $\varphi: \mathcal{A} \to \overline{\mathbb{R}}$ 为 $\mathcal{A}$ 上的 σ–**可加集函数**, 如果 $\varphi(\emptyset) = 0$ 且

$$\varphi(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \varphi(A_k)$$

对一切两两不交的 $A_k \in \mathcal{A}$, $k \in \mathbb{N}$, 且 $\cup_{k=1}^{\infty} \in \mathcal{A}$ 成立. σ–代数上的 σ–可加集函数也称为 **符号测度**, 若 φ 不以 $\pm\infty$ 为值, 则称 φ 为 **有限的**.

设 φ 是可测空间 $(\Omega, \mathcal{F})$ 上的符号测度， 于是对任何 $A \in \mathcal{F}$, $\varphi(A) + \varphi(A^c)$ 应该有意义， 因此如果有一 $A \in \mathcal{F}$ 使 $\varphi(A) = \pm\infty$,

则 $\varphi(\Omega)=\pm\infty$. 所以符号测度最多只能以 ∞, $-\infty$ 之一为值. 类似的论证可知：若 $\varphi(A)\in\mathbb{R}, A\in\mathcal{F}$, 则对任何 $B\subset A, B\in\mathcal{F}$, 都有 $\varphi(B)\in\mathbb{R}$.

显然非负符号测度是测度，而由推论 5.4.5 知当 $\int_\Omega f\,d\mu$ 存在时，f 在 $\mathcal{F}$ 上对 μ 的不定积分 $\varphi(A):=\int_A f\,d\mu$, $A\in\mathcal{F}$ 是符号测度. 再由

$$\varphi(A)=\int_A f^+ d\mu-\int_A f^- d\mu,\quad A\in\mathcal{F},$$

知不定积分可表为两个测度之差，自然可以问是否任一符号测度都可分解为两个测度之差？回答是肯定的. 我们现在来讨论这一问题.

对于符号测度，有类似于定理 3.3.2 及 3.3.3 的下列两条结论，其证明方法也完全类似. 留给读者自己证明.

2. 引理 设 φ 是 $(\Omega,\mathcal{F})$ 上的符号测度. 则 φ 下连续：即若 $\{A_n: n\in\mathbb{N}\}\subset\mathcal{F}$, $A_n\uparrow$, 则有

$$\lim_{n\to\infty}\varphi(A_n)=\varphi\Big(\bigcup_{n=1}^{\infty}A_n\Big);$$

且 φ 上连续：即若 $\{A_n: n\in\mathbb{N}\}\subset\mathcal{F}$, $A_n\downarrow$, 且有一 $m\in\mathbb{N}$ 使 $\varphi(A_m)\in\mathbb{R}$, 则有

$$\lim_{n\to\infty}\varphi(A_n)=\varphi\Big(\bigcap_{n=1}^{\infty}A_n\Big).\qquad\square$$

3. 引理 设 $(\Omega,\mathcal{F})$ 为可测空间，φ 为 $\mathcal{F}$ 上的有限可加集函数. 若 φ 满足下列两条件之一：

1) φ 下连续；

2) φ 有限且在 $\emptyset$ 处上连续 (即对任何 $\{A_n: n\in\mathbb{N}\}\subset\mathcal{F}$, $A_n\downarrow$ $\cap_{n=1}^{\infty}A_n=\emptyset$, 且存在 $m\in\mathbb{N}$, 使 $\varphi(A_m)\in\mathbb{R}$ 都有

$$\lim_{n\to\infty}\varphi(A_n)=\varphi(\emptyset)=0),$$

则 φ 为 $\mathcal{F}$ 上的符号测度.

4. 定理 设 φ 是 $(\Omega,\mathcal{F})$ 上的符号测度，则存在 $P,N\in\mathcal{F}$ 使

$$\varphi(P)=\sup_{A\in\mathcal{F}}\varphi(A),\qquad \varphi(N)=\inf_{A\in\mathcal{F}}\varphi(A),$$

即 φ 在 $\mathcal{F}$ 上达到最大值与最小值；且 $\Omega=P\cup N$, $P\cap N=\varnothing$.

证明 先考察对一切 $A\in\mathcal{F}$, $\varphi(A)<\infty$ 的情形.

这时，存在一集序列 $\{A_n:\ n\in\mathbb{N}\}\subset\mathcal{F}$ 使

$$\lim_{n\to\infty}\varphi(A_n)=\sup_{A\in\mathcal{F}}\varphi(A).$$

令 $A=\cup_{n=1}^{\infty}A_n$. 证明的思路是将 A 按下列方式无限分“细”，而 P 就是其中使 φ 取正值的那些集之并.

第一次对 A 的分划为 $A=A_1\cup(A\setminus A_1)$, 一般的第 n 次分划为

$$A=\bigcup_m A_{nm},\qquad A_{nm}:=\bigcap_{k=1}^{n}\tilde{A}_k,$$

其中 $\tilde{A}_k:=A_k$ 或 $A\setminus A_k$. 容易看出：每一次分划都是前一次分划的加密，所以当 $n'>n$ 时，每一 A_{nm} 都是若干个 $A_{n'm'}$ 的并.
令

$$B_n:=\bigcup_m A^+_{nm}.\qquad A^+_{nm}=\begin{cases}A_{nm}, & 当\ \varphi(A_{nm})>0,\\ \varnothing, & 当\ \varphi(A_{nm})\le 0.\end{cases}$$

即 B_n 为一切使 φ 取正值的 A_{nm}(n固定) 的并 (若这样的 A_{nm} 不存在，则令 $B_n=\varnothing$). 还有当 $N>n$ 时，每一 A^+_{Nm} 或为 B_n 的子集，或与 B_n 不相交，所以当 $N>n$ 时，

$$\begin{aligned}B_{nN}&:=B_n\cup\cdots\cup B_N\\ &=B_n\cup\Big(\bigcup_{A^+_{n+1,m}\cap B_n=\varnothing}A^+_{n+1,m}\Big)\cdots\cup\Big(\bigcup_{A^+_{Nm}\cap B_{n,N-1}=\varnothing}A^+_{Nm}\Big).\end{aligned}$$

后一表达式的集两两不交，所以可用可加性，因而 $\varphi(B_n)\le\varphi(B_{nN})$. 另一方面，由 $A_n=\cup(\tilde{A}_1\cap\cdots\cap\tilde{A}_{n-1}\cap A_n)$ 可见 A_n 是

一些 A_{nm} 的并，于是由 B_n 的定义知 $\varphi(A_n) \le \varphi(B_n)$. 故

$$\varphi(A_n) \le \varphi(B_{nN}).$$

令 $N \to \infty$. 由 φ 的下连续性 (引理 2) 即得

$$\varphi(A_n) \le \varphi(P_n), \qquad P_n := \bigcup_{k=n}^{\infty} B_k,$$

且易见 $P_n \downarrow P = \cap_{n=1}^{\infty} P_n$. 注意到 φ 不取 ∞ 为值及 $\varphi(P_n) \geqslant 0$, 因而 $\varphi(P_n) \in \mathbb{R}$, 利用 φ 的上连续性即得

$$\sup_{A\in\mathcal{F}} \varphi(A) = \lim_{n\to\infty} \varphi(A_n) \le \lim_{n\to\infty} \varphi(P_n) = \varphi(P).$$

这就证明了 P 的存在性. 再令 $N := P^c$, 则对任何 $A \in \mathcal{F}$ 都有 $\varphi(A^c) \le \varphi(P) < \infty$, 因而

$$\varphi(N) = \varphi(\Omega) - \varphi(P) \le \varphi(\Omega) - \varphi(A^c) = \varphi(A).$$

故在此情形下，$N \in \mathcal{F}$ 的存在性获证，且 $P \cup N = \Omega$, $P \cap N = \emptyset$.

对于 $\varphi(\Omega) = \infty$ 的情形，则由定义 1 后面的讨论知对一切 $A \in \mathcal{F}$, $\varphi(A) > -\infty$. 此时与前段证明类似 (或将前面的结论用于 $-\varphi$), 可证 $N \in \mathcal{F}$ 的存在, $P := N^c \in \mathcal{F}$, 因而 $P \cup N = \Omega$, $P \cap N = \emptyset$ 成立. □

5. 定理 (Hahn 分解定理) 设 φ 是 $(\Omega, \mathcal{F})$ 上的符号测度，则

1) 存在 $P, N \in \mathcal{F}$, $P \cup N = \Omega$, $P \cap N = \emptyset$, 使对一切 $A \in \mathcal{F}$ 有

$$\varphi(A \cap P) = \sup_{B\in\mathcal{F}\cap A} \varphi(B), \tag{1}$$

$$\varphi(A \cap N) = \inf_{B\in\mathcal{F}\cap A} \varphi(B), \tag{2}$$

其中 $\mathcal{F} \cap A := \{F \cap A : F \in \mathcal{F}\}$.

2) 若对任何 $A \in \mathcal{F}$, 令

$$\varphi^+(A) := \varphi(A \cap P),$$

$$\varphi^-(A) := -\varphi(A \cap N),$$

$$|\varphi|(A) := \varphi^+(A) + \varphi^-(A),$$

则 φ^+, φ^-, $|\varphi|$ 都是 $\mathcal{F}$ 上的测度, φ^+, φ^- 有一有限, 且

$$\varphi = \varphi^+ - \varphi^-.$$

φ^+、φ^-, $|\varphi|$ 分别称为 φ 的 **上、下、全变差**.

证明 1). 由于证法的类似, 只证 φ 不取 ∞ 为值的情形.

今往证: 定理 4 的证明中所取的 $P, N \in \mathcal{F}$ 符合要求. 由定理 4 的证明知

$$0 = \varphi(\emptyset) \leqslant \sup_{A \in \mathcal{F}} \varphi(A) = \varphi(P) < \infty.$$

由 (1) 立即可以看出: 对给定的 A 要证 (1), 只要证对一切 $B \in \mathcal{F} \cap A$ 有 $\varphi(A \cap P) \geq \varphi(B)$. 如若不然, 则存在 $B \in \mathcal{F} \cap A$ 使 $\varphi(A \cap P) < \varphi(B)$, 于是

$$\varphi(P) = \varphi(A \cap P) + \varphi(A^c \cap P) < \varphi(B \cup (A^c \cap P)),$$

而这与 P 的特性矛盾. 故 (1) 获证.

在此情形下 $N = P^c$. 对给定的 A, 若 $\varphi(A \cap N) = -\infty$, 则 (2) 式成立; 而当 $\varphi(A \cap N) > -\infty$ 时, 对一切 $B \in \mathcal{F} \cap A$, 则由 φ 不取 ∞ 为值及 (1) 得

$$\varphi(A \cap N) = \varphi(A) - \varphi(A \cap P) \leq \varphi(A) - \varphi(A \setminus B) = \varphi(B).$$

故 (2) 获证.

2). 由 (1),(2) 的右边知 φ^+, φ^- 非负, 而由 (1),(2) 的左边知 φ^+, φ^- 为 σ-可加的, 所以 φ^+, φ^-, $|\varphi|$ 都是 $\mathcal{F}$ 上的测度. 再由前边的证明知: 当 φ 不以 $\infty(-\infty)$ 为值时, $\varphi^+(\varphi^-)$ 有限. 由 φ^+, φ^- 的定义立知有最后的结论. □

定理 5 是对符号测度而言的. 如果 φ 只是集代数上的 σ–可加集函数, 则有

6. 定理 若 φ 是集代数上 $\mathcal{A}$ 的 σ–可加集函数, 则

$$\varphi = \varphi^+ - \varphi^-,$$

其中

$$\varphi^-(A) := -\inf_{B\in\mathcal{A}\cap A} \varphi(B), \quad A\in\mathcal{A},$$
$$\varphi^+(A) := \sup_{B\in\mathcal{A}\cap A} \varphi(B), \quad A\in\mathcal{A}.$$

且 φ^+, φ^- 以及 $|\varphi| := \varphi^+ + \varphi^-$ 都是 $\mathcal{A}$ 上的测度.

证明 参阅 [YWL] 第三章 §7.1 定理 3.

定理 5 和 6 的意义在于指出: σ–代数或集代数上的 σ–可加集函数都可以表成两个测度之差. 从而在许多情况下, 关于一般 σ–可加集函数的研究可以归结为测度的研究, 而测度比一般 σ–可加集函数更容易掌握.

在经典的实变函数论中, 与定理 5 、6 相当的是有界变差函数的分解, 现介绍如下:

7. 定义 设 F 为定义域包含 $[a,b]$,$a,b\in\mathbb{R}$, $a<b$ 的实数值函数, 令

$$\mathfrak{I} := \left\{\{t_k : 0\le k\le n\} :\ a=t_0<\cdots<t_n=b,\ n\in\mathbb{N}\right\},$$
$$V_F[a,b] := \sup\left\{\sum_k |F(t_k)-F(t_{k-1})| :\ \{t_k\}\in\mathfrak{I}\right\}.$$

若 $V_F[a,b]<\infty$, 则称 F 在 $[a,b]$ 上具 **有限变差**(或 **有界变差**). 称 $V_F[a,b]$ 为 F 在 $[a,b]$ 上的 **变差**. 类似地定义 F 在 $(-\infty,b]$ 上的

变差及 $\mathbb{R}$ 上的变差：

$$V_F(-\infty,b] := \sup\left\{\sum_k |F(t_k)-F(t_{k-1})| :\ \{t_k\}\in \mathcal{I}_{(-\infty,b]}\right\};$$

$$V_F(\mathbb{R}) := \sup\left\{\sum_k |F(t_k)-F(t_{k-1})| :\ \{t_k\}\in \mathcal{I}_{\mathbb{R}}\right\},$$

其中

$$\mathcal{I}_{(-\infty,b]} := \{\{t_k : 0\leqslant k\leqslant n\} :\ -\infty < t_0 < \cdots < t_n = b,\ n\in\mathbb{N}\},$$

$$\mathcal{I}_{\mathbb{R}} := \{\{t_k : 0\leqslant k\leqslant n\} :\ -\infty = t_0 < \cdots < t_n < \infty,\ n\in\mathbb{N}\}.$$

因而如果 $V_F(-\infty,b]$ ($V_F(\mathbb{R})$) 有限，称 F **在** $(-\infty,b]$ ($\mathbb{R}$) **上具有限变差**. 若 $F:\mathbb{R}\mapsto\mathbb{R}$ 具有限变差，则定义

$$V_F:\ \mathbb{R}\mapsto\mathbb{R},\quad V_F(x) := V_F(-\infty,x],\ x\in\mathbb{R},$$

设 μ 是 $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ 上的有限符号测度. 定义

$$F_\mu:\mathbb{R}\to\mathbb{R},\quad F_\mu(x) := \mu((-\infty,x]),\ x\in\mathbb{R}. \tag{1}$$

若 $\{t_k : 0\leq k\leq n\}\subset\mathbb{R}$, 且 $t_0<\cdots<t_n$, 则

$$\sum_{k=1}^n |F_\mu(t_k)-F_\mu(t_{k-1})| = \sum_{k=1}^n |\mu((t_{k-1},t_k])| \leq |\mu|(\mathbb{R}),$$

由此 $V_{F_\mu}(\mathbb{R})\leq|\mu|(\mathbb{R})$, 因而 F_μ 具有限变差. 并且由引理 2 易知它右连续且 $F_\mu(-\infty)=0$.

下面将证明 μ 与 F_μ 是一对一的. 为此先证明

8. 引理 设 F 是 $\mathbb{R}$ 上具有限变差的实值函数，则

1) F 有界且仅有第一类间断点；

2) $V_F[a,b]$ 有以下性质：

$$V_F(-\infty,b] = V_F(-\infty,a] + V_F[a,b],\quad a<b,\ a,b\in\mathbb{R}, \tag{1}$$

$$V_F(-\infty,b] = \lim_{a\to-\infty} V_F[a,b],\qquad b\in\mathbb{R}, \tag{2}$$

$$V_F[a,b] = \lim_{c\to a+} V_F[c,b],\quad \text{若} a<b \text{且} F(a+0)=F(a); \tag{3}$$

3) V_F 有界、不降且 $V_F(-\infty)=0$;

4) 若 F 右连续, 则 V_F 也右连续.

证明 1) F 有界显然. 往证 F 仅有第一类间断点. 只证对任何 $c\in\mathbb{R}, F(c+0)$ 存在. 否则, 若存在 $c\in\mathbb{R}$ 及 $a_n\downarrow c, b_n\downarrow c$, 使得 $\lim_{n\to\infty}F(a_n)=A, \lim_{n\to\infty}F(b_n)=B, A\neq B$. 不妨设对任何 $n\in\mathbb{N}$, $a_{n+1}<b_{n+1}<a_n<b_n$. 存在 n_0, 当 $n>n_0$ 时, $|F(b_n)-B|<\frac{|A-B|}{4}, |F(a_n)-A|<\frac{|A-B|}{4}$, 于是当 $N\to\infty$ 时,

$$V_F[c,b]\geqslant\sum_{n=n_0}^{n_0+N}|F(b_n)-F(a_n)|\geqslant N(|A-B|-\frac{|A-B|}{2})\to\infty.$$

2) 由定义 7 容易验证 (1). 为证 (2), 首先注意 $V_F[a,b]$ 是 a 的不增函数, 因而

$$\lim_{a\to-\infty}V_F[a,b]=\sup_{a<b}V_F[a,b].$$

另一方面,

$$V_F(-\infty,b]=\sup_{\substack{-\infty<t_0<t_1<\cdots<t_n=b\\ n\in\mathbb{N}}}\sum_{k=1}^{n}|F(t_k)-F(t_{k-1})|,$$

$$\sup_{t_0<b}V_F[t_0,b]=\lim_{a\to-\infty}V_F[a,b].$$

再利用 $V_F[-\infty,b]\geqslant V_F[a,b]$, 对一切 $a<b$ 因而 (2) 成立. 为证 (3), 只需证 $V_F[a,a+\varepsilon]\to 0$, 当 $\varepsilon\downarrow 0$. 事实上, 任给 $\varepsilon>0$, 存在 $\delta>0$, 当 $x,y\in[a,a+\delta]$ 时, $|F(x)-F(y)|<\frac{\varepsilon}{4}$, 另一方面, 存在 $a=t_0<t_1<t_2=a+\delta<t_3<\cdots<t_n=b$, 使 $V_F[a,b]<\sum_{k=1}^{n}|F(t_k)-F(t_{k-1})|+\frac{\varepsilon}{2}$, 于是

$$\begin{aligned}&V_F[a,a+\delta]+V_F[a+\delta,b]\\&\qquad<\sum_{k=3}^{n}|F(t_k)-F(t_{k-1})|+|F(a+\delta)-F(t_1)|\end{aligned}$$

$$\begin{aligned}&+ F(t_1) - F(a)| + \frac{\varepsilon}{2}\\&< V_F[a+\delta, b] + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{2} = V_F[a+\delta, b] + \varepsilon,\end{aligned}$$

即 $V_F[a, a+\delta] < \varepsilon$. (3) 获证.

3): 显然 V_F 有界且不降, 而由 (1),(2) 知

$$\begin{aligned}\lim_{x\to-\infty} V_F(-\infty, x] &= \lim_{x\to-\infty}(V_F(-\infty, b] - V_F[x, b])\\&= V_F(-\infty, b] - V_F(-\infty, b] = 0.\end{aligned}$$

故 3) 获证.

应用 (1),(3), 使用类似的方法可证 4). □

9. 定理 设 F 为 $\mathbb{R}$ 上具有限变差的函数. 则 F 可表为两个有界不降函数 F^+、F^- 之差: $F = F^+ - F^-$.

若 F 右连续且 $F(-\infty) = 0$, 则 F^+, F^- 亦然.

证明 容易验证函数

$$F^+ := \frac{V_F + F}{2}, \quad F^- := \frac{V_F - F}{2}$$

满足定理的要求. □

10. 定理 设 μ 为可测空间 $(\Omega, \mathcal{F})$ 上的有限符号测度, F_μ 由 (7.1) 定义. 则 $\mu \to F_\mu$ 决定 $(\Omega, \mathcal{F})$ 上的一切有限符号测度集与一切具有限变差右连续且在 $-\infty$ 处为零的函数集之间的一一映射.

证明 在定义 7 下面的讨论中, 已证 F_μ 右连续具有限变差且 $F_\mu(-\infty) = 0$. 若 μ 及 ν 是使 $F_\mu = F_\nu$ 的符号测度, 往证 $\mu = \nu$. 令

$$\mathcal{S} := \{(a, b] : a < b,\ a, b \in \bar{\mathbb{R}}\},$$

$$\Lambda := \{B \in \mathcal{B} : \mu(B) = \nu(B)\},$$

则对任何 $a, b \in \bar{\mathbb{R}}$, $a < b$, 有

$$\mu((a, b]) = F_\mu(b) - F_\mu(a) = F_\nu(b) - F_\nu(a) = \nu((a, b]),$$

于是 $\Lambda\supset\mathcal{S}$, 而 $\mathcal{S}$ 是 π–系, 由 μ, ν 具有 σ–可加性及下连续性易知 Λ 是一 λ–系. 故由集合形式的单调类定理知 $\Lambda\supset\sigma(\mathcal{S})=\mathcal{B}$, 因而 $\mu=\nu$.

今往证上述映射是满射, 即需证: 设 F 是 $\mathbb{R}$ 上的右连续、具有限变差且 $F_\mu(-\infty)=0$ 的函数, 则有 $\mathbb{R}$ 上的有限符号测度 μ 使得 $F=F_\mu$. 由定理 9, 有两个有界不降右连续且在 $-\infty$ 处为零的函数 F^+, F^- 使 $F=F^+-F^-$. 再由测度扩张定理的推论 3.2.18 知 F^+,F^- 分别决定 $\mathbb{R}$ 上的有限测度 μ_1, μ_2. 令 $\mu=\mu_1-\mu_2$, 则显然有 $F=F_\mu$, 故定理获证. □

习题

1. 设 φ, μ_1, μ_2 分别是可测空间 $(\Omega,\mathcal{F})$ 上的符号测度和测度, 且 $\varphi=\mu_1-\mu_2$. 则必有 $\varphi^+\le\mu_1$, $\varphi^-\le\mu_2$, 其中 φ^+, φ^- 如定理 5 所定义 (这叫做 Hahn 分解的最小性).

2. 设 φ, μ 分别是可测空间 $(\Omega,\mathcal{F})$ 上的有限可加集函数和有限测度. 若对任何 $A_n:\ n\in\mathbb{N}$, 当 $\mu(A_n)\to 0$ 时有 $\varphi(A_n)\to 0$, 则 φ 是符号测度.

3. 试证定理 9 中的

$$F^+(x):=\sup\left\{\sum_{k=1}^{n}[F(t_k)-F(t_{k-1}]^+:\ t_0<\cdots<t_n\le x,\ n\in\mathbb{N}\right\},$$

$$F^-(x):=\sup\left\{\sum_{k=1}^{n}[F(t_k)-F(t_{k-1}]^-:\ t_0<\cdots<t_n\le x,\ n\in\mathbb{N}\right\},$$

其中

$$[b-a]^+:=\begin{cases}b-a, & \text{当 } b\geqslant a,\\ 0, & \text{当 } b<a;\end{cases}\qquad [b-a]^-:=\begin{cases}0, & \text{当 } b\geqslant a,\\ a-b, & \text{当 } b<a.\end{cases}$$

4. 试证下列各函数在其定义域上具有限变差: (以下设 $a<b$, $a,b\in\mathbb{R}$)

1) $[a,b]$ 上的单调函数 F;

2) $[a,b]$ 上的 Lipschitz 函数 F(即, $F:[a,b]\to\mathbb{R}$, 且有一常数 K 使对一切 $x,y\in[a,b]$, $|F(x)-F(y)|\le K|x-y|$);

3) F 在 $[a,b]$ 上有有界导数.

5. 试证有限变差函数有界. 并举一反例说明逆命题不真.

6. 设 F, G 是 $[a,b]\subset\mathbb{R}$ 上的有限变差函数, 则它们的和差积仍然具有限变差; 若还有 $\inf_{a<x<b}|G(x)|>0$, 则 F/G 亦然.

§2. Lebesgue 分解定理与 Radon-Nikodym 定理

本节将讨论不定积分的刻划问题, 即给定测度空间 $(\Omega,\mathcal{F},\mu)$, 问 $\mathcal{F}$ 上的符号测度 φ 满足什麼条件时, 它是关于 μ 的不定积分? 这就是 Radon-Nikodym 定理. 为了其它附带的目的, 将证明更一般的 Lebesgue 分解定理.

先讨论不定积分应该具有的性质.

1. 定理 设 $(\Omega,\mathcal{F},\mu)$ 是测度空间, $\mathcal{F}$–可测函数 f 对 μ 的积分存在, 则不定积分 $\varphi(A):=\int_A f\,d\mu$, $A\in\mathcal{F}$ 存在且具有下列性质:

1) 若 $\mu(A)=0$, 则 $\varphi(A)=0$;

2) φ 是 $\mathcal{F}$ 上的符号测度;

3) 若 f a.e. 有限, μ 是 σ–有限测度, 则 φ 是 σ–有限集函数, 特别当 f 可积时, φ 有限.

证明 由积分性质易见 φ 存在及 1), 2) 成立. 今往证 3):

由于 μ 是 σ–有限测度, 存在 A_n, $n\in\mathbb{N}$ 两两不交, $\Omega=\cup_{n\in\mathbb{N}}A_n$ 且 $\mu(A_n)<\infty$, $n\in\mathbb{N}$. 其次若令 $B:=\{\omega:\ f(\omega)=\pm\infty\}$, $B_m:=\{\omega:\ m\le f(\omega)<m+1\}$, $m\in Z$ 则 B, B_m, $m\in Z$ 两两不

交且 $\Omega = B \cup (\cup_{m\in Z} B_m)$. 因而 B、$A_n \cap B_m$、$n \in \mathbb{N}$, $m \in Z$ 两两不交且

$$\Omega = B \cup (\cup_{m\in Z} \cup_{n\in\mathbb{N}} (A_n \cap B_m)).$$

其次由 f a.e. 有限知 $\mu(B) = 0$, 因此 $\varphi(B) = 0$; 又

$$\begin{aligned}-\infty < m\mu(A_n \cap B_m) &\le \varphi(A_n \cap B_m)\\ &= \int_{A_n\cap B_m} f\,d\mu \le (m+1)\mu(A_n \cap B_m) < \infty.\end{aligned}$$

故 φ σ–有限. □

不定积分的性质 1) 称为 μ–连续性, 一般地有下列的定义.

2. 定义 设 $(\Omega, \mathcal{F}, \mu)$ 为一测度空间, φ 是 $\mathcal{F}$ 上的一集函数. 若对任意的 $A \in \mathcal{F}$, $\mu(A) = 0$, 有 $\varphi(A) = 0$, 则称 φ 为 μ–**连续**, 记作 $\varphi \ll \mu$.

正如本节开头所说, 我们先证明更广的 Lebesgue 分解定理而将 Radon-Nikodym 定理作为它的推论. 为此先给出下列的

3. 定义 设 $(\Omega, \mathcal{F}, \mu)$ 是一测度空间, φ 是 $\mathcal{F}$ 上的一集函数. 若存在 $N \in \mathcal{F}, \mu(N) = 0$ 使得对所有的 $A \in \mathcal{F}$, 都有

$$\varphi(A \cap N^c) = 0.$$

其中 $N^c = \Omega \setminus N$, 则称 φ 为 μ–**奇异**.

显然, 若 μ, φ 都是测度, 则 φ 为 μ–奇异当且仅当 μ 为 φ–奇异.

4. 定理 (Lebesgue 分解定理). 设 $(\Omega, \mathcal{F}, \mu)$ 是一测度空间, φ 是 $\mathcal{F}$ 上的符号测度, 且 μ, φ 都是 σ–有限的, 则

$$\varphi = \varphi_c + \varphi_s, \tag{1}$$

其中 φ_c 是某一 $\mathcal{F}$–可测函数 f 的不定积分, 因而是 μ–连续的符号测度, 而 φ_s 是 μ–奇异的符号测度.

若 φ 不以 $-\infty$ 为值，则 φ_c、φ_s 亦然

上述分解唯一，而且 f 也是 μ a.e. 唯一决定，即若 g 也满足上述要求，则 $f=g$，μ-a.e..

(1) 称为 φ 的 Lebesgue **分解**.

证明 (A) 先证分解的存在性. 由符号测度的分解定理知任何符号测度均可表成两个测度的和而有限情形又是 σ–有限的基础，因而下面分三个步骤证明分解的存在性.

1) μ、φ 是有限测度；

2) μ、φ 是 σ–有限测度；

3) φ 是 σ–有限符号测度，μ 是 σ–有限测度.

先证 1) 设

$$\Phi := \left\{ f:\ f \geq 0,\ \text{可测};\ \int_A f\, d\mu \leq \varphi(A), \forall A \in \mathcal{F} \right\}.$$

因为 $0 \in \Phi$, 故 Φ 非空. 取一可测函数列 $\{f_n \in \Phi:\ n \in \mathbb{N}\}$ 使

$$\lim_{n\to\infty} \int f_n = \sup_{f\in\Phi} \int f =: \alpha \leq \varphi(\Omega) < \infty,$$

令 $g_n := \sup_{k\leq n} f_k$. 则 $0 \leq g_n \uparrow f := \sup_{n\in\mathbb{N}} f_n$. 又设

$$A_k := \{\omega:\ g_n(\omega) = f_k(\omega)\},$$

$$B_1 := A_1,\quad B_k := A_1^c \cap \cdots \cap A_{k-1}^c \cap A_k,\ k = 2, \cdots, n,$$

则 B_k, $k = 1, \cdots, n$ 两两不交且

$$\bigcup_{k=1}^{n} B_k = \bigcup_{k=1}^{n} A_k = \Omega.$$

对任何 $A \in \mathcal{F}$, 由单调收敛定理及 $B_k \subset A_k$ 得

$$\begin{aligned}\int_A f\, d\mu &= \lim_{n\to\infty} \sum_{k=1}^{n} \int_{A\cap B_k} f_k\, d\mu \\ &\leq \lim_{n\to\infty} \sum_{k=1}^{n} \varphi(A \cap B_k) = \varphi(A),\end{aligned}$$

类似地，有 $\int f\,d\mu=\alpha$. 故 $f\in\Phi$.

今往证：

$$0\le\varphi_c(A):=\int_A fd\mu,\quad A\in\mathcal{F},$$
$$0\le\varphi_s(A):=\varphi(A)-\varphi_c(A),\quad A\in\mathcal{F},$$

就是定理所要求的分解，为此只需证 φ_s 是 μ–奇异的.

令 $\varphi_n:=\varphi_s-\mu/n$, 显然对任何 $n\in\mathbb{N}$, φ_n 是有限符号测度，由 Hahn 分解定理知存在 $D_n\in\mathcal{F}$ 使对一切 $A\in\mathcal{F}$,

$$\varphi_n(A\cap D_n)\le 0,\quad \varphi_n(A\cap D_n^c)\ge 0.$$

取 $D:=\cap_1^\infty D_n$, 则对任意 $A\in\mathcal{F}$ 及 $n\in\mathbb{N}$,

$$\varphi_n(A\cap D)=\varphi_n(A\cap D\cap D_n)\le 0,$$

由 φ_n 的定义即得

$$0\le\varphi_s(A\cap D)\le\frac{1}{n}\mu(A\cap D).$$

令 $n\to\infty$, 注意到 μ 有限，即知对任意 $A\in\mathcal{F}$ 有

$$\varphi_s(A\cap D)=0.$$

于是为了证明 φ_s 的 μ–奇异性只需证明 $\mu(D^c)=0$, 而由于 $D^c=\cup_{n=1}^\infty D_n^c$, 所以只需证：

$$\mu(D_n^c)=0,\qquad \forall n\in\mathbb{N}.$$

事实上，对任何 $A\in\mathcal{F}$,

$$\int_A(f+\frac{1}{n}I_{D_n^c})\,d\mu=\varphi_c(A)+\frac{1}{n}\mu(A\cap D_n^c)$$
$$=\varphi(A)-\varphi_s(A)+\frac{1}{n}\mu(A\cap D_n^c).$$

注意到 $\varphi_s(A\cap D)=0$ 及 $D_n^c\subset D^c$, 上式

$$
\begin{aligned}
&= \varphi(A)-\varphi_s(A\cap D^c)+\frac{1}{n}\mu(A\cap D_n^c)\\
&\leq \varphi(A)-\varphi_s(A\cap D_n^c)+\frac{1}{n}\mu(A\cap D_n^c)\\
&= \varphi(A)-\varphi_n(A\cap D_n^c)\leq\varphi(A).
\end{aligned}
$$

这就是说 $f+(1/n)I_{D_n^c}\in\Phi$, 因此

$$
\alpha\geq\int_\Omega\left(f+\frac{1}{n}I_{D_n^c}\right)\,d\mu=\varphi_c(\Omega)+\frac{1}{n}\mu(D_n^c)=\alpha+\frac{1}{n}\mu(D_n^c).
$$

由于 $0\leq\alpha<\infty$, 故 $\mu(D_n^c)=0$. 因此当 φ,μ 都是有限测度的情形下，分解的存在性获证.

2) 现在证明：μ, φ 是 σ–有限测度时，分解的存在性. 此时可将 Ω 分划为 $\Omega=\cup_{n=1}^\infty A_n$, A_n, $n\in\mathbb{N}$ 两两不交且 $\varphi(A_n)$, $\mu(A_n)$, $n\in\mathbb{N}$ 有限. 于是 φ, μ 在 $\mathcal{F}\cap A_n$ 上的限制是 $\mathcal{F}\cap A_n$ 上的有限测度. 利用 1) 中的结论，对每一个 n, 存在两个有限测度 $\varphi_c^{(n)}$, $\varphi_s^{(n)}$、非负 $\mathcal{F}\cap A_n$–可测函数 f_n 及 $N_n\in\mathcal{F}\cap A_n$, 满足

$$
\begin{gathered}
\varphi(A_n\cap A)=\varphi_c^{(n)}(A_n\cap A)+\varphi_s^{(n)}(A_n\cap A),\\
0\leq\varphi_c^{(n)}(A_n\cap A)=\int_{A_n\cap A}f_n\,d\mu,\\
\varphi_s^{(n)}(A_n\cap A\cap N_n^c)=0,\ \mu(N_n)=0,\forall A\in\mathcal{F}.
\end{gathered}
$$

对每一 $A\in\mathcal{F}$, 令

$$
\begin{aligned}
\varphi_c(A)&:=\sum_{n=1}^\infty\varphi_c^{(n)}(A_n\cap A),\\
\varphi_s(A)&:=\sum_{n=1}^\infty\varphi_s^{(n)}(A_n\cap A).
\end{aligned}
$$

由上面的规定及已得结论易知： $\varphi(A)=\varphi_c(A)+\varphi_s(A)$, $A\in\mathcal{F}$, φ_c 是 $\mathcal{F}$–可测函数 $f:=\sum_{n=1}^\infty f_nI_{A_n}$ 关于 μ 的不定积分. 最后证

明：φ_s 是 μ–奇异的．令 $N := \cup_{n=1}^{\infty} N_n \in \mathcal{F}$，则显然有 $\mu(N) = 0$，而对于任意 $A \in \mathcal{F}$，有

$$\begin{aligned}\varphi_s(A \cap N^c) &= \sum_{n=1}^{\infty} \varphi_s^{(n)}(A_n \cap A \cap N^c) \\ &= \sum_{n=1}^{\infty} \varphi_s^{(n)}(A_n \cap A \cap N^c \cap N_n^c) = 0.\end{aligned}$$

这就证明了：当 φ, μ 都是 σ–有限测度时，分解的存在性．

还应注意：此时 φ_c、φ_s 都是测度，φ_c 是一非负可测函数的不定积分．

3) 最后证明：当 φ 是任意 σ–有限符号测度，μ 是 σ–有限测度时，分解的存在性．

由 Hahn 分解定理可知：

$$\varphi = \varphi^+ - \varphi^-,$$

其中 φ^+, φ^- 是两个测度且有一有限．不妨设 φ^- 有限．由 2) 的结论有

$$\varphi^+ = \varphi_c^+ + \varphi_s^+, \qquad \varphi^- = \varphi_c^- + \varphi_s^-,$$

其中 φ_c^+、φ_c^- 是关于 μ 的不定积分且 φ_c^- 有限；而 φ_s^+、φ_s^- 关于 μ 奇异且 φ_s^- 有限．令

$$\varphi_c := \varphi_c^+ - \varphi_c^-, \qquad \varphi_s := \varphi_s^+ - \varphi_s^-,$$

则

$$\varphi = \varphi_c + \varphi_s,$$

且 φ_c 是关于 μ 的不定积分（注意 φ_c^+、φ_c^- 有一有限），φ_s 是 μ 奇异的．

还有，若 φ 不以 $-\infty$ 为值，则 φ^- 有限，因而 φ_c^-，φ_s^- 有限，故 φ_c, φ_s 不以 $-\infty$ 为值．至此分解的存在性获证．

(B) 再证分解的唯一性. 设有两种分解:

$$\varphi = \varphi_c + \varphi_s = \varphi_c' + \varphi_s',$$

其中 φ_c, φ_c' 是 σ-可加、μ-连续的, 而 φ_s, φ_s' 是 σ-可加、μ-奇异的. 因此存在 $N_k \in \mathcal{F}$, $\mu(N_k) = 0$, $k = 1,2$ 使得对任何 $A \in \mathcal{F}$, 有

$$\varphi_s(A \cap N_1^c) = \varphi_s'(A \cap N_2^c) = 0.$$

令 $N = N_1 \cup N_2$, 则 $\mu(N) = 0$. 今往证对任何 $A \in \mathcal{F}$,

$$\varphi_c(A) = \varphi_c'(A), \quad \varphi_s(A) = \varphi_s'(A). \tag{1}$$

先证明 $\varphi(A)$ 有限的情形. 由假设

$$\varphi(A) = \varphi_c(A) + \varphi_s(A) = \varphi_c'(A) + \varphi_s'(A).$$

由于 $\varphi(A)$ 有限, 因此 $\varphi_c'(A)$, $\varphi_s(A)$ 有限. 于是

$$\varphi_c(A) - \varphi_c'(A) = \varphi_s'(A) - \varphi_s(A). \tag{2}$$

再由 $\varphi(A)$ 有限知 $\varphi(A \cap N^c)$ 有限, 于是由 (2) 有

$$\varphi_c(A \cap N^c) - \varphi_c'(A \cap N^c) = \varphi_s'(A \cap N^c) - \varphi_s(A \cap N^c).$$

利用 φ_c, φ_c' 的 μ-连续性, φ_s, φ_s' 的 μ-奇异性及它们的可加性, 即得

$$\begin{aligned}\varphi_c(A) - \varphi_c'(A) &= \varphi_c(A \cap N^c) - \varphi_c'(A \cap N^c) \\ &= \varphi_s'(A \cap N^c) - \varphi_s(A \cap N^c) \\ &= \varphi_s'(A \cap N_1^c \cap N_2^c) - \varphi_s(A \cap N_2^c \cap N_1^c) = 0.\end{aligned}$$

再利用 (2) 即得

$$\varphi_s'(A) - \varphi_s(A) = 0.$$

故对 $\varphi(A)$ 有限的情形 (1) 式获证.

再证 $\varphi(A)=\infty$ 的情形. 由于 φ 是 σ–有限, 故存在 $A_n\in\mathcal{F}$, $n\in\mathbb{N}$, 两两不交使 $\Omega=\cup_{n=1}^{\infty}A_n$, $\varphi(A_n)$ 有限, 于是由 σ–可加性及上述有限情形知对任何 $A\in\mathcal{F}$ 有

$$\varphi_c(A)=\cup_n\varphi_c(A\cap A_n)=\cup_n\varphi'_c(A\cap A_n)=\varphi'_c(A),$$
$$\varphi_s(A)=\cup_n\varphi_s(A\cap A_n)=\cup_n\varphi'_s(A\cap A_n)=\varphi'_s(A).$$

于是分解的唯一性得证. 至于 f 由 φ_ca.e.(μ) 决定是积分的性质. 至此, 定理全部获证. □

在 Lebesgue 分解定理中, 当 φ 具有 μ–连续性时, 就得知 φ 为某一函数的不定积分. 于是得到下列的重要结果.

5. 定理 (Radon-Nikodym 定理). 设 μ 是 Ω 中的 σ–代数 $\mathcal{F}$ 上的 σ–有限测度, φ 为 $\mathcal{F}$ 上的 σ–有限、μ–连续的符号测度. 则 φ 是某一有限 $\mathcal{F}$–可测函数 f 的不定积分, 且 f 由 φ 关于 μ 几乎唯一决定.

实际上, 这个结果还可以稍稍推广如下. 它在定义条件期望时有用.

5'. 定理 (Radon-Nikodym 定理的推广). 设 μ 是 Ω 中的 σ–代数 $\mathcal{F}$ 上的 σ–有限测度, φ 为 $\mathcal{F}$ 上的 μ–连续的符号测度 (未必 σ–有限). 则 φ 是某一 $\mathcal{F}$–可测函数 f(未必一定 a.e. 有限) 的不定积分, 且 f 由 φ 几乎唯一决定.

证明 . 我们只证明 φ 是测度且 μ 是有限测度的情形, 其它情形由 Hahn 分解定理及与定理 4 相同的手法得到. 请读者自己完成.

称 φ 在 A 上 σ–有限, 如果存在 $A_n\in\mathcal{F}$, $\varphi(A_n)<\infty$, $n\in\mathbb{N}$, 且 $\cup_nA_n=A$, 令

$$\mathfrak{B}:=\{A\in\mathcal{F}:\varphi\text{在}A\text{上}\sigma-\text{有限}\},$$

$$s := \sup_{B \in \mathfrak{B}} \mu(B).$$

则有一集序列 $B_n \in \mathfrak{B}$, $n \in \mathbb{N}$ 使 $s = \lim_{n\to\infty} \mu(B_n)$. 令

$$B := \bigcup_{n=1}^{\infty} B_n.$$

因 φ 在每一 B_n 上 σ–有限, 故 φ 在 B 上也 σ–有限, 因而 $B \in \mathfrak{B}$, 且 $s = \mu(B)$.

集类 $\mathcal{F}_1 := B \cap \mathcal{F}$ 和 $\mathcal{F}_2 := B^c \cap \mathcal{F}$ 分别是以 B 和 B^c 为空间的 σ–代数, φ 在 $\mathcal{F}_1$ 及 $\mathcal{F}_2$ 上的限制也分别是 $\mathcal{F}_1$ 及 $\mathcal{F}_2$ 上的 μ–连续测度, 且由 B 的定义知 φ 在 $\mathcal{F}_1$ 上 σ–有限. 故由定理 5 知有一定义在 B 上且 $\mathcal{F}_1$ 可测的函数 f_1 使得

$$\varphi(A) = \int_A f_1(\omega)\, d\mu, \quad A \in \mathcal{F}_1. \tag{1}$$

再考察 φ 在 $\mathcal{F}_2$ 上的情形. 由 μ–连续性显然有: 当 $\mu(A) = 0$, $A \in \mathcal{F}_2$ 时有 $\varphi(A) = 0$. 今往证:

$$\mu(A) > 0,\ A \in \mathcal{F}_2 \Longrightarrow \varphi(A) = \infty. \tag{2}$$

事实上, 若存在 $A \in \mathcal{F}_2$ 使 $\mu(A) > 0$ 而 $\varphi(A) < \infty$. 则一方面 φ 在 $A \cup B$ 上 σ–有限, 因而 $A \cup B \in \mathfrak{B}$, $\mu(A \cup B) \le s$, 另一方面, 由于 $A \cap B = \emptyset$ 有 $\mu(A \cup B) = \mu(A) + \mu(B) > s$. 这一矛盾证明了 (2).

由 (1),(2) 知: 若令

$$f(\omega) := \begin{cases} f_1(\omega), & \omega \in B \\ \infty, & \omega \in B^c. \end{cases}$$

则 φ 是 f 的不定积分. 至于 f 的 a.e. 唯一性由积分的性质容易推出. □

6. 定义 设 $(\Omega, \mathcal{F}, \mu)$ 是 σ–有限测度空间, φ 是 $\mathcal{F}$ 上 μ–连续的符号测度, 则按定理 5', μ–a.e. 唯一决定的函数 f 称为 φ 关于 μ 的 **Radon 导数**, 记作 $f = \dfrac{d\varphi}{d\mu}$.

由定理 5' 及定理 5 立得

7. **推论** 设 λ, μ 是 Ω 中 σ–代数 $\mathcal{F}$ 上的测度， μ 为 σ–有限，而 λ 为 μ–连续. 若 f 为 $\mathcal{F}$–可测函数，则 $\int f\,d\lambda$ 存在当且仅当 $\int f\frac{d\lambda}{d\mu}\,d\mu$ 存在，且当积分存在时，有

$$\int_A f\,d\lambda = \int_A f\frac{d\lambda}{d\mu}\,d\mu,\quad \forall A\in\mathcal{F}.$$

下面我们介绍 Lebesgue 分解定理对分布函数分解的应用. 为了简明起见，只介绍一元的情形.

8. **定理** 任意一维的有界分布函数 F 都可以分解为三个有界分布函数的和：

$$F = F_c + F_d + F_s,$$

其中 F_c 是 $\mathbb{R}$ 上的非负可积 Borel 函数 f 对 Lebesgue 测度 λ 的积分，即对一切 $x\in\mathbb{R}$,

$$F_c(x) = \int_{-\infty}^{x} f(t)\,dt, \tag{1}$$

F_c 称为 F 的 **绝对连续部分**; F_d 是一跳跃函数，即存在可数个实数对 (a_k, f_k), $f_k > 0$, $k\in\mathbb{N}$, 使

$$F_d(x) = \sum_{k:a_k\leqslant x} f_k,\qquad \forall x\in\mathbb{R},$$

F_d 称为 F 的 **离散部分**; F_s 是一连续分布函数，且它所对应的 L-S 测度对 L 测度奇异， F_s 称为 F 的 **奇异部分**.

证明 设 μ 是对应于 F 的 L-S 测度，由于 μ 与 F 的差分 ΔF 一一对应，所以自然考察 μ 的分解. 根据 Lebesgue 分解定理 (μ 看作那里的 φ, λ 看作那里的 μ)，存在分解式

$$\mu = \mu_c + \bar{\mu}_s,$$

其中 μ_c 是 λ-连续部分，可表成 $\mathbb{R}$ 上的非负有限 Borel 函数 f 对 λ 的积分. $\bar{\mu}_s$ 是 λ-奇异部分. 由假设知 μ_c, $\bar{\mu}_s$ 都是有限测度，因而 f 对 λ 可积.

为了得出另外两部分，将 $\bar{\mu}_s$ 再进行分解.

首先易证：使 $\bar{\mu}_s(\{a\}) > 0$ 的数 a 最多有可数个 (读者试自证之). 设它们是 a_k, $k \in \mathbb{N}$, 并令

$$f_k := \bar{\mu}_s(\{a_k\}), \qquad k \in \mathbb{N},$$

$$\mu_d(B) := \sum_{a_k \in B} f_k = \bar{\mu}_s(\{a_k : a_k \in B\}), \quad B \in \mathcal{B},$$

$$\mu_s := \bar{\mu}_s - \mu_d$$

显然 μ_d, μ_s 都是有限测度，且有

$$\mu = \mu_c + \mu_d + \mu_s.$$

令

$$F_c(x) := \mu_c(-\infty, x] = \int_{-\infty}^{x} f(t)\,dt,$$

$$F_d(x) := \mu_d((-\infty, x]) = \sum_{a_k \leq x} f_k,$$

$$F_s(x) := F(x) - F_d(x) - F_c(x), \quad \forall x \in \mathbb{R}.$$

则定理获证.

注 1). 定理 8 的结论从理论上说明：连续型概率分布与离散型概率分布及其非负线性组合是最常见的. 因此在较初等的教科书中，以它们为主要讨论对象是适宜的.

2). 对一般的 n 元分布函数有完全类似的结论，其证法也相同. 有兴趣的读者可参考 [YWL] 第三章 §7.3 定理 8.

9. 例 现在举一个由 Cantor 三分集出发构造的奇异分布函

数的例子. 让我们先复习 Cantor 集的概念，令

$$G_0 := \bigcup_{n\in\mathbb{N}} \bigcup_{a\in\{0,2\}^{n-1}} I_{(n,a)},$$

$$I_{(n,a)} := \Big(\frac{1}{3^n}\Big(\sum_{k=1}^{n-1} a_{n-k}3^k + 1\Big),\ \frac{1}{3^n}\Big(\sum_{k=1}^{n-1} a_{n-k}3^k + 2\Big)\Big),$$
$$(n,a) \in \mathbb{N}\times\{0,2\}^{n-1},$$

则 $P_0 = [0,1]\setminus G_0$ 就是 Cantor 集. 在 $D := (-\infty,0]\cup G_0\cup[1,\infty)$ 上定义函数 F 如下：

$$F(x) := \begin{cases} 0, & \text{当} x \le 0, \\ \frac{1}{2^n}\Big(\sum_{k=1}^{n-1}\frac{a_{n-k}}{2}2^k + 1\Big), & \text{当} x \in I_{(n,a)}, \\ 1, & \text{当} x \ge 1. \end{cases}$$

于是 F 在 D 上显然是不降的，在 G_0 的每一构成区间上为常数且

$$\lim_{x\downarrow 0} F(x) = 0, \quad \lim_{x\uparrow 1} F(x) = 1.$$

今再证 F 在 D 上一致连续. 因为 G_0 的前 $2^n - 1$ 个构成区间 $I_{(k,a)}$, $(k,a) \in \{k\}\times\{0,2\}^{k-1}$, $k = 1,\cdots,n$ 中任何两个之间的距离 $\ge 1/3^n$, 而 F 在从 $[0,1]$ 减去这些构成区间后所剩下的 2^n 个不相交的区间的每一个上的变差为 $1/2^n$, 所以有

$$0 \le y - x \le \frac{1}{3^n},\ x,y \in G_0, \implies 0 \le F(y) - F(x) \le \frac{1}{2^n}.$$

由此即知 F 在 D 上一致连续.

由 §3.2.3(VI), 习题 13 及上面已证事实知，由

$$\widetilde{F}(x) := \inf_{x<t\in D} F(t)$$

定义的 $\widetilde{F}$ 是 $\mathbb{R}$ 上的连续概率分布函数，且在 D 上与 F 一致. 再由于 F 在 D 的每一构成区间上为常数， $\widetilde{F}$ 所决定的概率测度 μ

在 D 上为 0, 即 $\mu(D)=0$. 另一方面 $D^c=P_0$ 的 Lebesgue 测度 $\lambda(P_0)=0$. 故 $\widetilde{F}$ 为一奇异分布函数. □

习题

在下面各题中，除特殊声明外， φ 表示可测空间 $(\Omega,\mathcal{F})$ 上符号测度 (2,6 除外), $\varphi=\varphi^+-\varphi^-$ 表示 Hahn 分解， $\bar{\varphi}:=\varphi^++\varphi^-$, μ 表示测度 $A,B,\cdots$ 表示 $\mathcal{F}$–可测集.

1. 若 $\mu(A_n)\to 0\ (n\to\infty)$, $A_n\in\mathcal{F}$ 时 $\varphi(A_n)\to 0\ (n\to\infty)$, 则 φ 是 μ 连续的；若 φ 是有限的，则反之亦真. (提示：首先，由 φ 的 μ 连续性易知 $|\varphi|$ 也是 μ 连续的. 若其逆不真，则存在 $\varepsilon>0$ 与序列 $A_n\in\mathcal{F}$, $n\in\mathbb{N}$ 使得 $\mu(A_n)<1/2^n$ 而 $|\varphi(A_n)|\geq\varepsilon$, 于是 $B:=\overline{\lim}_{n\to\infty}A_n$ 使得 $\mu(B)=0$ 而 $|\varphi|(B)\geq\varepsilon$.)

若 φ 是一般符号测度时如何？

2. 若 $\{\mu_n\}$ 是测度序列，试证：对任何 $n\in\mathbb{N}$,

$$\mu_n\ll\mu:=\sum_{k=1}^{\infty}\frac{1}{2^k}\mu_k.$$

μ_n 能否换成 φ_n ?

3. R.-N. 导数的微分公式：设 $\varphi\ll\nu$, 且 ν, φ, $\varphi'\ll\mu$, 则

$$\frac{d(\varphi+\varphi')}{d\mu}=\frac{d\varphi'}{d\mu}+\frac{d\varphi'}{d\mu}\quad \mu-\text{a.e.},$$

$$\frac{d\varphi}{d\mu}=\frac{d\varphi}{d\nu}\frac{d\nu}{d\mu}\quad \mu-\text{a.e.}.$$

4. 设 $\bar{\mu}_n:=\sum_{k=1}^n\mu_k\to\bar{\mu}$, $\bar{\nu}_n:=\sum_{k=1}^n\nu_k\to\bar{\nu}$, 其中带有附标的 μ, ν 都是有限的，并且 $\bar{\nu}_n\ll\bar{\mu}_n$, 对一切 $n\in\mathbb{N}$. 则

1) $\dfrac{d\mu_1}{d\bar{\mu}_n}\to\dfrac{d\mu_1}{d\bar{\mu}}$ $\bar{\mu}$–a.e.,

2) 若 $\bar{\mu}_n\ll\bar{\nu}$, 对一切 $n\in\mathbb{N}$, 则 $\dfrac{d\bar{\mu}_n}{d\bar{\nu}}\to\dfrac{d\bar{\mu}}{d\bar{\nu}}$ $\bar{\nu}$–a.e.,

3) $\bar{\nu} \ll \bar{\mu}$ 且 $\dfrac{d\bar{\nu}_n}{d\bar{\mu}_n} \to \dfrac{d\bar{\nu}}{d\bar{\mu}}$ $\bar{\mu}$–a.e..

(提示：关于最后一个结论应注意：若对一切 $n \in \mathbb{N}$, $\bar{\mu}_n(A_n) = 0$, 则 $\bar{\mu}(\overline{\lim}_{n\to\infty} A_n) = 0$. 由此可知，只要考察一个特别选择的

$$\frac{d\bar{\nu}_n}{d\bar{\mu}_n} = \sum_{k=1}^{n} f_k \Big/ \sum_{k=1}^{n} g_k,$$

其中 $f_k := \dfrac{d\nu_k}{d\bar{\mu}}, g_k := \dfrac{d\mu_k}{d\bar{\mu}}$, 而 $\sum_n f_n = \dfrac{d\bar{\nu}}{d\bar{\mu}}$, $\sum_n g_n = 1$ $\bar{\mu}$–a.e..)

5. 试证：要想 $\mathcal{F}$ 上的集函数 φ 是某一 $\mathcal{F}$ 可测函数 f 对 μ 的不定积分，必须且只需 φ 为 σ–可加，且对于每一集 $A := \{a \le f \le b\} \cap B$, $B \in \mathcal{F}$, 总有

$$a\mu(A) \le \varphi(A) \le b\mu(A).$$

6. 设 f 对 μ 的积分存在. 令 φ 是 f 关于 μ 的不定积分. 试证：

$$\varphi^+(A) = \int_A f^+ \, d\mu; \quad \varphi^-(A) = \int_A f^- \, d\mu;$$

$$|\varphi|(A) = \int_A |f| \, d\mu.$$

7. 设 f 为 $\mathcal{F}$ 可测函数，且使得下式右边有意义，则定义

$$\int f \, d\varphi := \int f \, d\varphi^+ - \int f \, d\varphi^-.$$

称为 **f 对 φ 的积分**. 试证：这种积分具有可测函数对测度的积分的主要性质.

8. 集代数 $\mathcal{A}$ 上的 σ–可加集函数 φ 可以扩张为 $\sigma(\mathcal{A})$ 上的符号测度的充分与必要条件是 φ 有下界或有上界. 若 φ 为 σ–有限，则扩张唯一且 σ–有限.

9. 若 μ 不是 σ–有限的，即使 φ 有限，则 Radon-Nikodym 定理也不一定成立.

提示:考虑反例: $\Omega := [0,1], \mathcal{F} := \{A: A \subset [0,1], A$或$A^c$可数, $\mu(A) := |A|(A$ 的元数).

$$\varphi(A) := \begin{cases} 0, & A\text{可数}, \\ 1, & A^c\text{可数}. \end{cases}$$

§3. 条件期望的概念

在概率论中, 条件期望是一个十分重要的概念. 本节的目的就是介绍它的一般概念及性质. 由于比较抽象, 我们将从初等概念出发引导出给出一般概念的途径, 以便于理解和掌握.

1. 我们知道事件 A 在给定事件 $B(\mathbf{P}(B) > 0)$ 的条件下的条件概率

$$\mathbf{P}(A|B) := \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)},$$

因而随机变量 X 在随机变量 Y 取值于 $(y, y + \Delta y]$ 的条件下的条件分布函数是

$$F_X(x|y < Y \leq y + \Delta y) := \frac{\mathbf{P}(X \leq x,\ y < Y \leq y + \Delta y)}{\mathbf{P}(y < Y \leq y + \Delta y)},$$

进而随机变量 X 在随机变量 Y 取值于 $(y, y + \Delta y]$ 的条件下的条件期望是

$$\mathbf{E}[X|y < Y \leq y + \Delta y] := \frac{\mathbf{E}[XI_{\{y<Y\leq y+\Delta y\}}]}{\mathbf{P}(y < Y \leq y + \Delta y)}.$$

于是随机变量 X 在随机变量 $Y = y$ 的条件下的条件分布函数与条件期望自然应该分别是

$$(1) \qquad F_X(x|Y = y) = \lim_{\Delta y \to 0} \frac{\mathbf{P}(X \leq x,\ y < Y \leq y + \Delta y)}{\mathbf{P}(y < Y \leq y + \Delta y)},$$

$$(2) \qquad \mathbf{E}[X|Y = y] = \lim_{\Delta y \to 0} \frac{\mathbf{E}[XI_{\{y<Y\leq y+\Delta y\}}]}{\mathbf{P}(y < Y \leq y + \Delta y)}.$$

由于 (1),(2) 两式的分母经常在 Δy 尚未趋於零时就为零，因此很难将它们作为定义. 但是将它们与导数定义比较，并注意到事件类 $\{\{y<Y\leq y+\Delta y\}:\ y\in\mathbb{R},\ \Delta y\in(0,\infty)\}$ 生成 σ–代数 $\sigma(Y)$，则它们分别相当于测度 $\mathbf{P}(\{X\leq x\}\cap B)$, $B\in\sigma(Y)$ 及 $\mathbf{E}[XI_B]$, $B\in\sigma(Y)$ 对 $\mathbf{P}_{\sigma(Y)}$(表示 $\mathbf{P}$ 在 $\sigma(Y)$ 上的限制) 的导数. 这就使我们想到它可以用 Radon 导数来定义. 从而引导出下列的定义. 为了减少叙述的头绪，先只给出条件期望的定义.

2. 定义 设 $(\Omega,\mathcal{F},\mathbf{P})$ 是概率空间，$\mathcal{C}$ 是 $\mathcal{F}$ 的子 σ–代数，X 的数学期望存在. 记 $\mathbf{P}$ 在 $\mathcal{C}$ 上的限制为 $\mathbf{P}_{\mathcal{C}}$. 则 $\mathbf{E}[XI_B]$, $B\in\mathcal{C}$, 对 $\mathbf{P}_{\mathcal{C}}$ 的 Radon 导数称为 **X 在 σ–代数 $\mathcal{C}$ 下 (关于 $\mathbf{P}$) 的条件期望**, 记作 $\mathbf{E}[X|\mathcal{C}]$.

因为 $\mathbf{E}[X]$ 存在，$\mathbf{E}[XI_B]$, $B\in\mathcal{C}$, 是 $\mathbf{P}_{\mathcal{C}}$ 上的符号测度，且 $\mathbf{P}_{\mathcal{C}}$ 连续，故由 Radon-Nikodym 定理的推广 (定理 2.5') 知：$\mathbf{E}[X|\mathcal{C}]$a.e. 唯一存在. 若 $\mathcal{C}=\sigma(Y)$, Y 为随机变量，则由定理 4.2.9 知，有一 Borel 可测函数 g 使

$$\mathbf{E}[X|\sigma(Y)]=g(Y)\ a.e.,$$

将它称为 **X 在 Y 之下的条件期望**, 并记作 $\mathbf{E}[X|Y]$. 于是自然可以将

$$\mathbf{E}[X|Y=y]:=g(y), \tag{1}$$

认为是 (1.2) 的确切表达.

由于 $\mathbf{P}(A)=\mathbf{E}[I_A]$, 我们称 $\mathbf{E}[I_A|\mathcal{C}]$ ($\mathbf{E}[I_A|Y]$) 为 **事件 $A\in\mathcal{F}$ 在 $\mathcal{C}$ (Y) 下的条件概率**.

下面给出几个可由定义直接得出的性质：

3. 命题 条件期望具有下列性质：

(I) 若 $\mathbf{E}[X]$ 存在，则 $\mathbf{E}[\mathbf{E}[X|\mathcal{C}]]$ 存在，且对任何 $B\in\mathcal{C}$,

$$\mathbf{E}[XI_B]=\int_B\mathbf{E}[X|\mathcal{C}]d\mathbf{P}_{\mathcal{C}}=\int_B\mathbf{E}[X|\mathcal{C}]d\mathbf{P},$$

因而 $\mathbf{E}[\mathbf{E}[X|\mathcal{C}]] = \mathbf{E}X$;

由此还得到：若 X 可积，则 $\mathbf{E}[X|\mathcal{C}]$ $\mathbf{P}$-a.e. 有限；

(II) 若 $\mathcal{C} = \mathcal{F}$ 或 X 为 $\mathcal{C}$–可测时，则 $\mathbf{E}[X|\mathcal{C}] = X, \mathbf{P}_{\mathcal{C}}$–a.e..

(III) 若 X 是 $\mathbf{E}X$ 存在的复随机变量， $X = X_1 + iX_2$, 则

$$\begin{aligned}\mathbf{E}[X|\mathcal{C}] &= \mathbf{E}[X_1|\mathcal{C}] + i\mathbf{E}[X_2|\mathcal{C}] \\ &= \mathbf{E}[X_1^+|\mathcal{C}] - \mathbf{E}[X_1^-|\mathcal{C}] + i(\mathbf{E}[X_2^+|\mathcal{C}] - \mathbf{E}[X_2^-|\mathcal{C}].\end{aligned}$$

证明 (I) 的第一式即定义，令 $B = \Omega$ 即得后一式. (II),(III) 由定义 2 及积分的线性性质即得. □

4. 命题 条件概率具有下列性质：

(I') 对任何 $A \in \mathcal{F}$, $B \in \mathcal{C}$, 有

$$\mathbf{P}(A \cap B) = \int_B \mathbf{P}(A|\mathcal{C})\, d\mathbf{P},$$

因而 $\mathbf{E}[\mathbf{P}(A|\mathcal{C})] = \mathbf{P}(A)$.

(II') 若 $A \in \mathcal{C}$, 则 $\mathbf{P}(A|\mathcal{C}) = I_A, \mathbf{P}_{\mathcal{C}}$–a.e..

条件期望与条件概率还有很多其它重要的性质，将在下节及下章予以讨论. 我们在此提醒读者注意：条件期望和条件概率是概率论中两个非常重要的概念，但是它们又比较抽象，不容易掌握. 为了更好地理解它，我们先讨论一些特殊的情形以及与初等概念的联系，然后再作几点重要的说明. 希望这样做能对读者有所帮助.

5. 定理 采用定义 2 的记号，则在 $\mathcal{C}$ 的每一非空原子 (即除 $\varnothing$ 及本身外不包含其它 $\mathcal{C}$–可测子集) B 上， $\mathbf{E}[X|\mathcal{C}]$ 是常数. 若还有 $\mathbf{P}(B) > 0$, 则

$$\mathbf{E}[X|\mathcal{C}] = \frac{1}{\mathbf{P}(B)} \int_B X\, d\mathbf{P}, \quad \forall \omega \in B. \tag{1}$$

因此记此常数为 $\mathbf{E}[X|B]$.

证明 由于 B 非空，故有一 $\omega\in B$. 则由 $\mathbf{E}[X|\mathcal{C}]$ 为 $\mathcal{C}$–可测函数知

$$B_0:=\{\omega:\ \mathbf{E}[X|\mathcal{C}](\omega)=\mathbf{E}[X|\mathcal{C}](\omega_0)\}\cap B\in\mathcal{C},$$

再由 $\omega\in B_0\subset B$ 及原子的定义知 $B_0=B$, 因而 $\mathbf{E}[X|\mathcal{C}]$ 在 B 上是常数，记为 $\mathbf{E}[X|B]$. 则由条件期望的定义知

$$\mathbf{E}[X|B]\mathbf{P}(B)=\int_B\mathbf{E}[X|\mathcal{C}]d\mathbf{P}=\int_B X\,d\mathbf{P}.$$

由 $\mathbf{P}(B)>0$ 即得另一结论. □

这个定理说明 $\mathbf{E}[X|\mathcal{C}]$ 在 $\mathcal{C}$ 的原子上是 X 的平均值，在这个意义上，$\mathbf{E}[X|\mathcal{C}]$ 是 X 的一个 $\mathcal{C}$–平滑函数. 由这条定理可以将定义 2 与初等概念以如下方式联系起来.

6. 推论 设 $\{B_n:\ n\in\mathbb{N}\}\subset\mathcal{F}$, 两两不交且 $\Omega=\cup_{n=1}^{\infty}B_n$(即 B_n, $n\in\mathbb{N}$ 是 Ω 的一个可数划分), $\mathcal{C}:=\sigma(B_n:\ n\in\mathbb{N})$, 则对任何期望存在的随机变量 X 来说,

$$\mathbf{E}[X|\mathcal{C}]=\sum_n\mathbf{E}[X|B_n]I_{B_n},$$

其中当 $\mathbf{P}(B_n)>0$ 时，$\mathbf{E}[X|B_n]$ 由 (3.1) 的右边给出，而当 $\mathbf{P}(B_n)$ =0 时，$\mathbf{E}[X|B_n]$ 取任一常数.

特别，对任何 $B\in\mathcal{F}$, $\mathbf{P}(B)>0$, 令 $\mathcal{C}:=\{\emptyset,B,B_n,\Omega\}$, 则对任何期望存在的随机变量 X 及随机事件 A 来说,

$$\mathbf{E}[X|B]=\frac{1}{\mathbf{P}(B)}\int_B X\,d\mathbf{P},\quad \mathbf{P}(A|B)=\frac{\mathbf{P}(A\cap B)}{\mathbf{P}(B)}.$$

这与初等概念一致.

最后，当 $\mathcal{C}:=\{\emptyset,\Omega\}$ 时，则有

$$\mathbf{E}[X|\mathcal{C}]=\mathbf{E}X.$$

证明 注意 B_n, B 都是 $\mathcal{C}$ 的原子，再应用定理 3 即得. □

7. 例 设 N 是一取非负整数值的随机变量, $\mathbf{E}N$ 有限; X_n, $n \in \mathbb{N}$ 是数学期望有限, 独立同分布的随机变量列, 且与 N 独立. 试证:

$$E := \mathbf{E}[\sum_{k=1}^{N} X_k] = \mathbf{E}X_1\mathbf{E}N. \tag{1}$$

这个例子在 6.3.10 中讨论过, 现在再用条件期望的概念加以讨论.

在解答这个问题之前, 先介绍一下它的实际原型:

问题 1. 求在单位时间内某一平面区域 G 上接受来自宇宙粒子的平均能量. 由于落入 G 的粒子数 N 是随机的, 所以可设它是一个取非负整数值的随机变量. 而粒子的能量也带有随机性, 自然可设为独立同分布的随机变量 X_k. 于是就将问题化为求 E.

问题 2. 设对一母体进行简单抽样, 但抽样是随机终止的. 于是求样本和的均值问题也化为求 E.

证明 令 $\mathcal{C} := \sigma(N) = \sigma(N = n:\ n \in \mathbb{N})$ 满足推论 6 的条件, 于是由推论 6 及命题 3(I) 即得

$$\begin{aligned} E &:= \mathbf{E}[\mathbf{E}[\sum_{k=1}^{N} X_k|\sigma(N)]] \\ &= \mathbf{E}[\sum_{n=1}^{\infty} \mathbf{E}[\sum_{k=1}^{N} X_k|N = n]I_{N=n}], \end{aligned} \tag{2}$$

其中 (并应用 N 与 $\{X_k,\ k \in \mathbb{N}\}$ 独立)

$$\begin{aligned} \mathbf{E}[\sum_{k=1}^{N} X_k|N = n] &= \frac{\mathbf{E}[I_{N=n}\sum_{k=1}^{N} X_k]}{\mathbf{P}(N = n)} \\ &= \mathbf{E}[\sum_{k=1}^{n} X_k] = n\mathbf{E}X_1. \end{aligned} \tag{3}$$

将 (3) 代入 (2) 即得

$$E = \mathbf{E}[X_1]\sum_{n=1}^{\infty} n\mathbf{E}[I_{N=n}] = \mathbf{E}X_1\mathbf{E}N.$$

8. 例 设 X,Y 是二维连续型随机变量，其分布密度为 $p(x,y)$, $(x,y)\in\mathbb{R}^2$, 求 X 在 Y 之下的条件期望 $\mathbf{E}(X|Y)$.

由定义 2 知存在一个 $\mathbb{R}$ 上的 Borel 可测函数 $g(y), y\in\mathbb{R}$, 使得 $\mathbf{E}(X|Y)=g(Y)$, $\mathbf{P}_{\sigma(Y)}$–a.e.. 再由条件期望的定义有

$$\int_{Y\in B} X\,d\mathbf{P}=\int_{Y\in B} g(Y)\,d\mathbf{P}\quad ,\forall\ B\in\mathcal{B}. \tag{1}$$

由积分变换定理及 Fibini 定理知：对一切 $B\in\mathcal{B}$,

$$\begin{aligned}&\int_{Y\in B} X\,d\mathbf{P}=\int_{\Omega} XI_B(Y)\,d\mathbf{P}\\ =&\int_{\mathbb{R}^2} xI_B(y)p(x,y))\,dxdy=\int_B\Big(\int_{\infty}^{\infty} xp(x,y)\,dx\Big)dy;\\ &\int_{Y\in B} g(Y)\,d\mathbf{P}=\int_{\Omega} g(Y)I_B(Y)\,d\mathbf{P}\\ =&\int_{\mathbb{R}^2} g(y)I_B(y)p(x,y)\,dxdy=\int_B g(y)\Big(\int_{-\infty}^{\infty} p(x,y)\,dx\Big)dy.\end{aligned}$$

因此，我们有

$$\int_{\infty}^{\infty} xp(x,y)\,dx=g(y)\Big(\int_{-\infty}^{\infty} p(x,y)\,dx\Big),\quad \lambda-\text{a.e.},$$

其中 λ 为 Lebesgue 测度. 若令 $N=\{y:\int_{-\infty}^{\infty}p(x,y)\,dx=0\}$, 则 $\mathbf{P}(Y\in N)=0$, 故而可取

$$g(y)=\begin{cases}\dfrac{\int_{-\infty}^{\infty} xp(x,y)\,dx}{\int_{-\infty}^{\infty} p(x,y)\,dx}, & \text{若 } \int_{-\infty}^{\infty} p(x,y)\,dx\neq 0\\ c, & \text{若 } \int_{-\infty}^{\infty} p(x,y)\,dx=0,\end{cases}$$

其中 c 为任意确定的常数. 显然 $g(Y)$ 是 $\sigma(Y)$ 可测的且满足 (1), 因而

$$\mathbf{E}(X|Y)=g(Y),\quad \mathbf{P}_{\sigma(Y)}-\text{a.e.}.$$

9. 注 关于条件期望和条件概率的几点重要说明：

1) 在定义中，即令当 X 取有限值时，它在 $\mathcal{F}$ 上关于 $\mathbf{P}$ 的不定积分是 σ–有限的，但未必在子 σ–代数 $\mathcal{C}$ 上 σ–有限，例如 $\mathcal{C} = \{\emptyset, \Omega\}$, $\varphi(\Omega) = \infty$. 因此必须用定理 2.5' 才能保证 $\mathbf{E}[X|\mathcal{C}]$ 的存在唯一性，并且 $\mathbf{E}[X|\mathcal{C}]$ 未必 a.e. 有限. 但当 X 可积时，$\mathbf{E}[X|\mathcal{C}]$ 也可积.

2) 在这一节及以后谈到条件期望与条件概率时，均指关于一个给定的概率 $\mathbf{P}$ 而言的. 为了叙述简单，我们一般省去'关于 $\mathbf{P}$'的定语.

3) 由于 $\mathbf{E}[X|\mathcal{C}]$ 只有在 $\mathbf{E}X$ 存在时才有意义，因此今后一概假定出现在条件期望符号下的随机变量 X 的数学期望 (即积分) $\mathbf{E}X$ 存在. 而且为了叙述简单，我们常常在证明有关条件期望的结论时，只讨论 X 有限或可积的情形.

4) 在谈到 $\mathbf{E}[X|\mathcal{C}]$ 的 a.e. 性质时，零概率集均指 $\mathcal{C}$–可测集而言，即存在 $\mathcal{C}$–可测集 N, $\mathbf{P}(N) = 0$, 使当 $\omega \in N_c$ 时，该性质成立. 若将 $\mathbf{P}$ 在 $\mathcal{C}$ 上的限制记成 $\mathbf{P}_{\mathcal{C}}$, 则可表成 $\mathbf{P}_{\mathcal{C}}$ a.e. 成立.

5) 给定概率空间 $(\Omega, \mathcal{F}, \mathbf{P})$, $\mathcal{C}$ 是 $\mathcal{F}$ 的子 σ–代数，令 $\mathfrak{E}$ 表示数学期望存在且 $\mathbf{P}$ a.e. 相等 $\mathcal{F}$–可测函数的等价类的集，而 $\mathfrak{G}$ 表示 $\mathbf{P}_{\mathcal{C}}$ a.e. 相等 $\mathcal{C}$–可测函数的等价类的集. 则符号 $\mathbf{E}[\bullet|\mathcal{C}]$ 是 $\mathfrak{E}$ 到 $\mathfrak{G}$ 的映射，我们称它为**给定 σ–代数 $\mathcal{C}$ 下的条件期望**. 还有也可以将 $\mathbf{E}[\bullet|\mathcal{C}](\bullet)$ 看成是由 $\Omega \times \mathfrak{E}$ 到 $\bar{\mathbb{R}} = [-\infty, \infty]$ 的函数，而在 (ω, X) 处的值就是 $\mathbf{E}[X|\mathcal{C}](\omega)$, 但是要特别注意此时是先取定 $\mathbf{E}[X|\mathcal{C}]$(它是 $\mathbf{P}_{\mathcal{C}}$ a.e. 唯一确定的), 然后才能唯一地给出 $\mathbf{E}[X|\mathcal{C}](\omega)$.

习题

1. 设随机变量 $X \sim \mathcal{P}(\lambda)$ (意指 X 服从具参数 λ 的 Poisson 分布), 随机变量 Y 在给定事件 $X = n$ 下的条件分布为

$$\mathbf{P}(Y = m|X = n) = \binom{n}{m} p^m (1-p)^{n-m}, \ m = 0, \cdots, n.$$

试证：　$Y \sim \mathcal{P}(p\lambda)$.

2. 用户在单位时间内向电话局要求通话的总时间的平均值称为该电话局的话务量. 设单位时间内用户向电话局呼唤的次数 $N \sim \mathcal{P}(\lambda)$ (表示 N 服从具参数 λ 的 Poisson 分布), 而每一用户的通话时间 $X_k \sim \mathcal{E}(\beta)$, (表示 X_k 服从具参数 β 的指数分布), $k \in \mathbb{N}$, 则该电话局的话务量是 λ/β.

3. 设二维随机变量 (X, Y) 具有概率分布密度 $p(x, y)$, 且 X 是可积随机变量. 则 $\mathbf{E}[X|X+Y=z]$ 可由下式给出：

$$\frac{\int xp(x, z-x)\,dx}{\int p(x, z-x)\,dx}.$$

4. 设二维随机变量 (X, Y) 的概率分布密度

$$p(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\{-\frac{1}{2(1-\rho^2)}(\frac{x^2}{\sigma_1^2} - 2\rho\frac{xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2})\},$$

其中 $\sigma_1 > 0$, $\sigma_2 > 0$, $\rho \in (0,1)$. 试求 $\mathbf{E}[Y|X=x]$.

§4. 条件期望的性质

由于条件概率是条件期望的特殊情形，所以今后主要讨论条件期望的性质. 而条件期望的性质主要有两部分，一部分与数学期望作为积分的性质相当，通常称之为期望性质；另一部分则与数学期望为“取平均值”的意义相当，通常称之为平滑性质. 下面我们首先讨论基本的期望性质.

1. **定理**　条件期望有下列的期望性质：

(I) 若 $X = a$, $\mathbf{P}$–a.e., 则 $\mathbf{E}[X|\mathcal{C}] = a$, $\mathbf{P}_{\mathcal{C}}$–a.e.;

(II) 若 $a\mathbf{E}X + b\mathbf{E}Y$ 存在, 则

$$\mathbf{E}[aX + bY|\mathcal{C}] = a\mathbf{E}[X|\mathcal{C}] + b\mathbf{E}[Y|\mathcal{C}],\ \mathbf{P}_{\mathcal{C}} - \text{a.e.};$$

(III) 若 $X \le Y$, $\mathbf{P}$–a.e., 则 $\mathbf{E}[X|\mathcal{C}] \le \mathbf{E}[Y|\mathcal{C}]$, $\mathbf{P}_{\mathcal{C}}$–a.e.;

(IV) (单调收敛性) 若 $0 \leq X_n \uparrow X$, $\mathbf{P}$–a.e., 则

$$0 \leq \mathbf{E}[X_n|\mathcal{C}] \uparrow \mathbf{E}[X|\mathcal{C}],\ \mathbf{P}_{\mathcal{C}} - \text{a.e.};$$

(V) (Fatou 收敛性) 设 Y, Z 可积，若对任何 $n \in \mathbb{N}, Y \leqslant X_n$, $\mathbf{P}$–a.e., 则

$$\mathbf{E}[\underline{\lim}_{n\to\infty} X_n|\mathcal{C}] \leq \underline{\lim}_{n\to\infty} \mathbf{E}[X_n|\mathcal{C}],\ \mathbf{P}_{\mathcal{C}} - \text{a.e.};$$

若对任何 $n \in \mathbb{N}, X_n \leq Z$, 则

$$\overline{\lim}_{n\to\infty} \mathbf{E}[X_n|\mathcal{C}] \leq \mathbf{E}[\overline{\lim}_{n\to\infty} X_n|\mathcal{C}],\ \mathbf{P}_{\mathcal{C}} - \text{a.e.};;$$

(VI) (控制收敛性). 设 Y, Z 可积，若 $Y \leq X_n \uparrow X$, $\mathbf{P}$–a.e., 或对任何 $n \in \mathbb{N}, Y \leq X_n \leq Z$ 且 $X_n \to X$, $\mathbf{P}$–a.e., 则

$$\mathbf{E}[X_n|\mathcal{C}] \longrightarrow \mathbf{E}[X|\mathcal{C}],\ \mathbf{P}_{\mathcal{C}} - \text{a.e.};$$

证明 由定义 3.2, (I) 显然成立. 其次由积分的线性性质知 $\mathbf{E}[aX + bY]$ 存在，再加上定义 3.2 可推出：对任何 $B \in \mathcal{C}$ 有

$$\begin{aligned}
\int_B \mathbf{E}[aX + bY|\mathcal{C}]\, d\mathbf{P} &= \int_B (aX + bY)\, d\mathbf{P} \\
&= a\int_B X\, d\mathbf{P} + b\int_B Y\, d\mathbf{P} \\
&= a\int_B \mathbf{E}[X|\mathcal{C}] d\mathbf{P} + b\int_B \mathbf{E}[Y|\mathcal{C}] d\mathbf{P} \\
&= \int_B [a\mathbf{E}[X|\mathcal{C}] + b\mathbf{E}[Y|\mathcal{C}]]\, d\mathbf{P}.
\end{aligned}$$

(注意上式中各项的存在性都是由 $\mathbf{E}X$, $\mathbf{E}Y$, $a\mathbf{E}X + b\mathbf{E}Y$ 的存在性所保证的). 上式说明了 $\mathbf{E}[aX + bY|\mathcal{C}]$ 及 $a\mathbf{E}[X|\mathcal{C}] + b\mathbf{E}[Y|\mathcal{C}]$ 在 $\mathcal{C}$ 上的不定积分相等，因而由推广的 R.-N. 定理知 (II) 成立.

再次，由积分的性质及定义 3.2 知，对任何 $B \in \mathcal{C}$ 有

$$\int_B \mathbf{E}[X|\mathcal{C}] d\mathbf{P} = \int_B X\, d\mathbf{P} \leq \int_B Y\, d\mathbf{P} = \int_B \mathbf{E}[Y|\mathcal{C}] d\mathbf{P}.$$

因而由积分的性质 $\mathbf{E}[X|\mathcal{C}] \leq \mathbf{E}[Y|\mathcal{C}]$, $\mathbf{P}$–a.e.(作为一个很好的复习，读者可以想想这里用到那些积分的性质). 故 (III) 获证.

(IV) 的证明: 由 (III) 知 $\mathbf{E}[X_n|\mathcal{C}]$ 是 $\mathbf{P}_{\mathcal{C}}$–a.e. 不降且非负的, 因此有一 $\mathcal{C}$–可测函数 X' 使

$$0 \le \mathbf{E}[X_n|\mathcal{C}] \uparrow X', \ \mathbf{P}_{\mathcal{C}} - \text{a.e.},$$

再由积分的单调收敛定理知, 对任何 $B \in \mathcal{C}$ 有

$$\lim_{n\to\infty} \int_B \mathbf{E}[X_n|\mathcal{C}]d\mathbf{P} = \int_B X'd\mathbf{P}.$$

另一方面由定义 3.2 及单调收敛定理知, 对任何 $B \in \mathcal{C}$ 有

$$\begin{aligned}\lim_{n\to\infty} \int_B \mathbf{E}[X_n|\mathcal{C}]d\mathbf{P} &= \lim_{n\to\infty} \int_B X_n d\mathbf{P} \\ &= \int_B X d\mathbf{P} = \int_B \mathbf{E}[X|\mathcal{C}]d\mathbf{P}.\end{aligned}$$

因而有

$$\int_B X'd\mathbf{P} = \int_B \mathbf{E}[X|\mathcal{C}]d\mathbf{P}, \quad \forall B \in \mathcal{C}.$$

由于 X' 与 $\mathbf{E}[X|\mathcal{C}]$ 都是 $\mathcal{C}$–可测函数, 故 $X' = \mathbf{E}[X|\mathcal{C}]$, $\mathbf{P}_{\mathcal{C}}$–a.e., 因而 (IV) 获证.

(V) 的证明与 Fatou 定理的证明是一样的, 只需将应用单调收敛定理之处相应地换成 (IV) 即可 (留给读者作为习题). (VI) 是 (IV)、(V) 的直接推论. □

条件期望具有与数学期望相应的一系列不等式 (在下一章讨论), 也可以认为是期望性质. 但是证明时要用到平滑性质, 并且平滑性质本身在理论上有很重要的作用, 所以我们先讨论条件期望的平滑性质. 首先有下列简单的性质.

2. 定理 若 $\mathcal{C}$ 与 $\sigma(X)$ 独立, 则

$$\mathbf{E}[X|\mathcal{C}] = \mathbf{E}X, \ \mathbf{P}_{\mathcal{C}} - \text{a.e.}.$$

证明 由假设知对任何 $B\in\mathcal{C}$ 来说， I_B 与 X 独立，因而由命题 3.3(I) 知

$$\mathbf{E}[I_B\mathbf{E}[X|\mathcal{C}]]=\mathbf{E}[I_BX]=\mathbf{P}(B)\mathbf{E}X=\mathbf{E}[I_B\mathbf{E}X].$$

故定理获证. □

由命题 3.3(II) 知当 X 是 $\mathcal{C}$–可测函数时，

$$\mathbf{E}[X|\mathcal{C}]=X,\ \mathbf{P}_{\mathcal{C}}-\text{a.e.}.$$

根据前面的解释，可以认为 $\mathbf{E}[\cdot|\mathcal{C}]$ 对 $\mathcal{C}$–可测函数不再起平滑作用. 更一般地有下列的

3. 定理 若 X 是 $\mathcal{C}$–可测的随机变量， Y 是使 $\mathbf{E}[XY]$、$\mathbf{E}Y$ 存在的随机变量 ($\mathcal{F}$–可测), 且 X、Y 之一是实的，或 XY,Y 是可积的复随机变量，则

$$\mathbf{E}[XY|\mathcal{C}]=X\mathbf{E}[Y|\mathcal{C}],\ \mathbf{P}_{\mathcal{C}}-\text{a.e.}.$$

证明 为简单起见只讨论 XY,Y 都是可积的情形. 当 X 为简单函数 I_C, $C\in\mathcal{C}$ 时，对任何 $B\in\mathcal{C}$ 有

$$\begin{aligned}\int_B\mathbf{E}[I_CY|\mathcal{C}]d\mathbf{P}&=\int_B I_CY\,d\mathbf{P}\\&=\int_{B\cap C}\mathbf{E}[Y|\mathcal{C}]d\mathbf{P}=\int_B I_C\mathbf{E}[Y|\mathcal{C}]d\mathbf{P}.\end{aligned}$$

故定理对此情形成立. 因而定理对 X 是 $\mathcal{C}$–简单函数也成立. 对于 XY,Y 都可积的情形，则由下列引理立刻得到.

引理 若 X 是 $\mathcal{C}$–可测函数，则存在 $\mathcal{C}$–简单函数列 X_n, $n\in\mathbb{N}$ 使 (i) $|X_n|\le|X|$, (ii) $X_n\to X$.

在证明引理之前我们先应用引理证明定理的一般情形. 事实上, 取引理的 $\mathcal{C}$–简单函数列 X_n, $n\in\mathbb{N}$, 则 $|X_nY|\le|XY|$, $X_nY\to XY$, 因而 X_nY 可积. 于是由定理对 $\mathcal{C}$–简单函数情形及条件期望的控制收敛性即知定理对一般情形成立.

现在我们再来证明引理：当 X 为非负函数时

$$X_n := \sum_{k=0}^{n2^n-1} \frac{k}{2^n} I_{\{k/2^n \le X < (k+1)/2^n\}},\ n \in \mathbb{N},$$

即符合要求. 当 X 为实函数时，则存在两个 $\mathcal{C}$–简单函数列 X_{in}, $n \in \mathbb{N}$, $i = 1, 2$, 使 $0 \le X_{1n} \uparrow X^+$、$0 \le\uparrow X^-$. 于是 $X_n := X_{1n} - X_{2n}$ 即符合要求. 最后当 $X = X^1 + iX^2$ 为复可测函数时, 则存在 $\mathcal{C}$–简单函数列 $X_n^i, n \in \mathbb{N}, i = 1, 2$ 使 $|X_n^i| \le |X^i|$, $X_n^i \to X^i$, $i = 1, 2$. 于是 $X_n := X_n^1 + iX_n^2$, $n \in \mathbb{N}$ 即符合要求. 至此引理获证. 因而定理对 XY、Y 可积的情形完全证明. □

4. 定理 若 $\mathcal{C} \subset \mathcal{C}'$, 则

$$\mathbf{E}[\mathbf{E}[X|\mathcal{C}]|\mathcal{C}'] = \mathbf{E}[X|\mathcal{C}] = \mathbf{E}[\mathbf{E}[X|\mathcal{C}']|\mathcal{C}],\ \mathbf{P}_{\mathcal{C}} - \text{a.e.}.$$

证明 由于 $\mathbf{E}[X|\mathcal{C}]$ 为 $\mathcal{C}$–可测函数，因而由 $\mathcal{C} \subset \mathcal{C}'$ 知它是 $\mathcal{C}'$–可测的，再由 $\mathbf{E}[\mathbf{E}[X|\mathcal{C}]] = \mathbf{E}X$ 存在及定理 3 知

$$\mathbf{E}[\mathbf{E}[X|\mathcal{C}]|\mathcal{C}'] = \mathbf{E}[X|\mathcal{C}]\mathbf{E}[1|\mathcal{C}'] = \mathbf{E}[X|\mathcal{C}],\ \mathbf{P}_{\mathcal{C}} - \text{a.e.},$$

此即第一个结论. 至于第二个结论，则由于对任何 $B \in \mathcal{C}$ (因而 $B \in \mathcal{C}'$) 来说，由定义 3.2 知

$$\int_B \mathbf{E}[\mathbf{E}[X|\mathcal{C}']|\mathcal{C}]d\mathbf{P} = \int_B \mathbf{E}[X|\mathcal{C}']d\mathbf{P} = \int_B X\, d\mathbf{P} = \int_B \mathbf{E}[X|\mathcal{C}]d\mathbf{P}.$$

再由不定积分的性质即得第二结论.

5. 定理 条件期望具有性质：若 X 可积，则

$$|\mathbf{E}[X|\mathcal{C}]| \le \mathbf{E}[|X||\mathcal{C}],\ \mathbf{P}_{\mathcal{C}} - \text{a.e.}.$$

证明 由假设 X 可积，令

$$\mathbf{E}[X|\mathcal{C}](\omega) = r(\omega)e^{i\theta(\omega)},$$

其中当 $r(\omega)=0$ 时，取 $e^{i\theta(\omega)}=1$. 并设 $X=Y+iZ$、Y,Z 是实随机变量. 则由 (II) 知

$$r=|\mathbf{E}[X|\mathcal{C}]|=\sqrt{(\mathbf{E}[Y|\mathcal{C}])^2+(\mathbf{E}[Z|\mathcal{C}])^2}$$

是 $\mathcal{C}$–可测函数. 于是由

$$e^{-i\theta(\omega)}=\begin{cases}\frac{r(\omega)}{\mathbf{E}[X|\mathcal{C}](\omega)}, & r(\omega)\neq 0,\\ 1, & r(\omega)=0,\end{cases}$$

知 $e^{-i\theta}$ 也是 $\mathcal{C}$–可测函数. 由条件期望的平滑性质 (定理 3)

$$|\mathbf{E}[X|\mathcal{C}]|=\mathbf{E}[X|\mathcal{C}]e^{-i\theta}=\mathbf{E}[Xe^{-i\theta}|\mathcal{C}]\geq 0.$$

故对任何 $B\in\mathcal{C}$,

$$\begin{aligned}0\leq\int_B|\mathbf{E}[X|\mathcal{C}]|d\mathbf{P}&=\int_B Xe^{-i\theta}\,d\mathbf{P}=\left|\int_B Xe^{-i\theta}\,d\mathbf{P}\right|\\ &\leq\int_B|X|\,d\mathbf{P}=\int_B\mathbf{E}[|X||\mathcal{C}]d\mathbf{P}.\end{aligned}$$

□

习题

1. 在条件期望的控制收敛性 (VI) 中，其它条件均成立，将 $X_n\to X$, $\mathbf{P}$–a.e. 换成 $X_n\xrightarrow{\mathbf{P}}X$, 结果如何？

2. 1) 若 $\mathcal{C}$, $\mathcal{C}'$ 是 $\mathcal{F}$ 的子 σ–代数， $\mathcal{C}\subset\mathcal{C}'$ 且 X' 是 $\mathcal{C}'$ 可测的， $\mathbf{E}XX'$, $\mathbf{E}X$ 有限，则

$$\mathbf{E}[XX'|\mathcal{C}]=\mathbf{E}[X'\mathbf{E}[X|\mathcal{C}']|\mathcal{C}],\ \mathbf{P}_{\mathcal{C}}-\text{a.e.}.$$

2) 若 1) 成立，则定理 3.4 成立.

3. 设 $\mathcal{F}_n$, $n\in\mathbb{N}$ 是 $\mathcal{F}$ 的子 σ–代数列，且 $\mathcal{F}_n\uparrow$, X_n 是 $\mathcal{F}_n$–可测的， $n\in\mathbb{N}$. 若对任何 n 及 $m>n$, 有

$$\mathbf{E}[X_m|\mathcal{F}_n]=X_n\ (\geq X_n,\ \leq X_n),\ \text{a.e.},$$

则称 $\{X_n: n\in\mathbb{N}\}$ 为一 $\mathcal{F}_n$ 鞅 (相应的:下鞅, 上鞅), $\sigma(X_1,\cdots,X_n)$ 鞅 (下鞅, 上鞅) 简称为鞅 (相应地: 下鞅, 上鞅). 试证明:

1) $\{X_n: n\in\mathbb{N}\}$ 为 $\mathcal{F}_n$ 鞅 (下鞅, 上鞅) 的充分与必要条件是

$$\mathbf{E}[X_{n+1}|\mathcal{F}_n]=X_n(相应地:\geq X_n,\ \leq X_n);$$

2) 设 $Y_n, n\in\mathbb{N}$ 为独立随机变量序列, 若对任何 $n\in\mathbb{N}$, 有 $\mathbf{E}Y_n=0$ (相应的: $\geq 0,\leq 0$), 则

$$\{X_n:=\sum_{k=1}^{n}Y_k:\ n\in\mathbb{N}\}$$

为鞅 (相应地: 下鞅, 上鞅).

*4. 随机变量序列 $\{X_n: n\in\mathbb{Z}_+\}$ 称为 **马尔可夫过程**, 如果对任何 $n\in\mathbb{Z}_+$ 及 $B\in\mathcal{B}$, 有

$$\mathbf{P}(X_{n+1}\in B|X_1,\cdots,X_n)=\mathbf{P}(X_{n+1}\in B|X_n)\text{a.e.}. \tag{1}$$

试证: (1) 与下列各命题等价 (都是指对任何 $n\in\mathbb{Z}_+$ 而言):

$$\mathbf{E}[Y|X_1\cdots,X_n]=\mathbf{E}[Y|X_n],\quad \forall Y\in\sigma(X_{n+1}). \tag{2}$$

$$\mathbf{E}[Y_1\cdots Y_m|X_1,\cdots,X_n]=\mathbf{E}[Y_1\cdots Y_m|X_n],\quad \forall Y_k\in\sigma(X_{n+k}),k=1,\cdots,m. \tag{3}$$

$$\mathbf{E}[Y|X_1,\cdots,X_n]=\mathbf{E}[Y|X_n],\quad \forall Y\in\sigma(\{X_m:m>n\}). \tag{4}$$

$$\mathbf{P}(FB|X_n)=\mathbf{P}(F|X_n)\mathbf{P}(B|X_n),\quad \forall B\in\sigma(X_1,\cdots,X_n),F\in\sigma(\{X_m:m>n\}). \tag{5}$$

5. 设随机变量 X,Y 满足

$$\mathbf{E}[Y^2|\mathcal{C}]=X^2\text{ a.e.},\quad \mathbf{E}[Y|\mathcal{C}]=X,\ \text{a.e.},$$

则 $X=Y$ a.e..

*§5. 条件概率分布

为了叙述简单起见，我们用 $(\Omega,\mathcal{F},\mathbf{P})$ 表示任意固定的概率空间. 注意到 §3 所定义的条件概率与条件数学期望具有与概率、数学期望非常类似的性质，即：

1°. 函数 $\mathbf{P}(A|\mathcal{C})$ 具有性质

$$\mathbf{P}(\Omega|\mathcal{C})=1 \text{ a.e.};\qquad \mathbf{P}(A|\mathcal{C})\geqslant 0 \text{ a.e.};$$

$$\mathbf{P}(\cup_{k=1}^{\infty}A_k|\mathcal{C})=\textstyle\sum_{k=1}^{\infty}\mathbf{P}(A_k|\mathcal{C}) \text{ a.e.},$$

其中 $A_k, k\in\mathbb{N}$, 两两不交.

2°. 函数 $E(X|\mathcal{C})$ a.e. 具有 X 关于 $\mathbf{P}(\cdot|\mathcal{C})$ 积分的性质，即

$$E(\textstyle\sum_{k=1}^{n}a_kI_{A_k}|\mathcal{C})=\sum_{k=1}^{n}a_k\mathbf{P}(I_{A_k}|\mathcal{C}); \text{ a.e.}$$

$$0\leqslant X_n\uparrow X\Rightarrow E(X_n|\mathcal{C})\uparrow E(X|\mathcal{C}), \text{a.e.}$$

但是，并不能由 1°, 2° 直接得出 $E(X|\mathcal{C})$ 是 X 关于 $\mathbf{P}(A|\mathcal{C})$, $A\in\mathcal{F}$ 的积分. 因为 $E(X|\mathcal{C})$ 及 $\mathbf{P}(A|\mathcal{C})$ 都是 ω 的函数，所以谈到 $E(X|\mathcal{C})$ 是 X 关于 $\mathbf{P}(A|\mathcal{C})$, $A\in\mathcal{F}$ 的积分的含义时应该是指：对于每一个 $\omega\in\Omega$ 或至少存在一个与 A 无关的零概率集 N, 使得对每一 $\omega\in N^c$, $E(X|\mathcal{C})(\omega)$ 是 X 关于 $\mathbf{P}(\cdot|\mathcal{C})(\omega)$ 的积分. 这就首先要求对每个 $\omega\in N^c$, $\mathbf{P}(A|\mathcal{C})(\omega)$, $A\in\sigma(X)$ 是 $\sigma(X)$ 上的概率. 但是，1° 中诸性质中的例外集既与给定的集 $A, A_k, k=1,2,\cdots$ 有关，又与条件概率的取法有关，而 $\sigma(X)$ 中元未必只有可数多个，因此是否存在共同的例外零概率集不是显而易见的.

为了解决上述问题，我们引进下面的

1. **定义** 设 $\mathcal{F}_1,\mathcal{C}$ 是 $\mathcal{F}$ 的子 σ–域， $\Omega\times\mathcal{F}_1$ 上函数 $\mathbf{P}^{\mathcal{C}}(\omega,A)$ 若满足

1) 对一切 $\omega \in \Omega$, $\mathbf{P}^{\mathcal{C}}(\omega, \cdot)$ 是 $\mathcal{F}_1$ 上概率测度；

2) 对一切 $A \in \mathcal{F}_1$, $\mathbf{P}^{\mathcal{C}}(\cdot, A)$ 是 $\mathcal{C}$ 可测函数，且存在 $\mathcal{C}$ 可测的零概率集 N, 使得

$$\mathbf{P}^{\mathcal{C}}(\omega, A) = \mathbf{P}(A|\mathcal{C})(\omega), \ \omega \in N^c.$$

则称 $\mathbf{P}^{\mathcal{C}}(\omega, A)$, $(\omega, A) \in \Omega \times \mathcal{F}_1$ 为 $\mathcal{F}_1$ **在 $\mathcal{C}$ 之下的正则条件概率**.

2. 例 设 $\{A_n, n \in \mathbb{N}\}$ 是 Ω 的可测分割，$\mathcal{C} = \sigma(\{A_n, n \in \mathbb{N}\})$. 则

$$\mathbf{P}^{\mathcal{C}}(\omega, A) = \begin{cases} \frac{\mathbf{P}(A \cap A_n)}{\mathbf{P}(A_n)}, & \omega \in A_n \text{ 且 } \mathbf{P}(A_n) > 0; \\ \mathbf{P}(A), & \omega \in A_n \text{ 且 } \mathbf{P}(A_n) = 0, \end{cases} \qquad A \in \mathcal{F},$$

为 $\mathcal{F}$ 在 $\mathcal{C}$ 之下的正则条件概率.

3. 定理 设 $\mathbf{P}^{\mathcal{C}}(\omega, A)$, $(\omega, A) \in \Omega \times \mathcal{F}_1$ 为 $\mathcal{F}_1$ 上在 $\mathcal{C}$ 之下的正则条件概率，X 为 $\mathcal{F}_1$ 可测函数，则对任何一个取定的条件期望 $E(X|\mathcal{C})$, 存在一个 $\mathbf{P}_{\mathcal{C}}$ 零概率集 N 使得

$$E(X|\mathcal{C})(\omega) = \int_\Omega X(\omega')\mathbf{P}^{\mathcal{C}}(\omega, d\omega'), \qquad \omega \in N^c. \tag{1}$$

证明 我们只需对一切非负 $\mathcal{F}_1$ 可测函数证明 (1) 成立. 首先，若 $X = I_A$，$A \in \mathcal{F}_1$, 由定义立得；其次，若 $X = \sum_{k=1}^n a_k I_{A_k}$，$a_k \geqslant 0$，$A_k \in \mathcal{F}_1$，$k = 1, \cdots, n$. 则由条件期望的性质知，对取定的 $E(X|\mathcal{C})$, $\mathbf{P}(A_k|\mathcal{C})$, $k = 1, \cdots, n$, 存在 $\{N_0, N_1, \cdots, N_n\} \subset \mathcal{C}$, $\mathbf{P}(N_k) = 0$, $k = 0, 1, \cdots, n$, 使得

$$E(X|\mathcal{C})(\omega) = \sum_{k=1}^n a_k \mathbf{P}(A_k|\mathcal{C})(\omega), \quad \omega \in N_0^c,$$

$$\mathbf{P}(A_k|\mathcal{C})(\omega) = \mathbf{P}^{\mathcal{C}}(\omega, A_k)(\omega), \quad \omega \in N_k^c, k = 1, \cdots, n.$$

令 $N = \bigcup_{k=0}^n N_k$. 则

$$\left.\begin{aligned} E(X|\mathcal{C})(\omega) &= \sum_{k=1}^n a_k \mathbf{P}(A_k|\mathcal{C})(\omega), \\ \sum_{k=1}^n \mathbf{P}^{\mathcal{C}}(\omega, A_k)(\omega) &= \int_\Omega X(\omega')\mathbf{P}^{\mathcal{C}}(\omega, d\omega') \end{aligned}\right\} \quad \omega \in N^c,$$

于是 (1) 对非负 $\mathcal{F}_1$–简单函数成立. 若 X 为 $\mathcal{F}_1$ 非负可测函数, 则存在 $\mathcal{F}_1$ 非负简单函数列 X_n, $n \in \mathbb{N}$ 及 $\mathcal{C}$–可测零概率集列 N_k, $k = 0, 1, 2, \cdots$, 使得 $X_n \uparrow X$ 且

$$E(X|\mathcal{C})(\omega) = \lim_{n \to \infty} E(X_n|\mathcal{C})(\omega), \quad \omega \in N_0^c.$$

$$E(X_n|\mathcal{C})(\omega) = \int_\Omega X_n(\omega')\mathbf{P}^{\mathcal{C}}(\omega, d\omega'), \quad \omega \in N_n^c, n \in \mathbb{N}.$$

令 $N = \cup_{k=0}^{\infty}$, 则 (1) 成立. □

但是, 正则条件概率未必总是存在的, 反例见 [Wg],p135. 我们进而给出

4. 定义 若 X 为 $(\Omega, \mathcal{F}, \mathbf{P})$ 到距离可测空间 $(E, \mathcal{B})$ 上的可测映射, $\mathcal{C}$ 为 $\mathcal{F}$ 的子 σ–域, $\Omega \times \mathcal{B}$ 上的函数 $\mathbf{P}_X^{\mathcal{C}}(\omega, B)$ 若满足

1) 对每个 ω, $\mathbf{P}_X^{\mathcal{C}}(\omega, \cdot)$ 是 $\mathcal{B}$ 上的概率测度;

2) 对每个 $B \in \mathcal{B}$, $\mathbf{P}_X^{\mathcal{C}}(\cdot, B)$ 是 $\mathcal{C}$ 可测函数且

$$\mathbf{P}_X^{\mathcal{C}}(\cdot, B) = \mathbf{P}(X^{-1}(B)|\mathcal{C})(\cdot), \quad \mathbf{P}_{\mathcal{C}} - \text{a.e.}$$

则称 $\mathbf{P}_X^{\mathcal{C}}(\omega, B), (\omega, B) \in \Omega \times \mathcal{B}$ 为 **X 在 $\mathcal{C}$ 之下的条件概率分布**.

与定理 3 类似可以证明下述定理.

5. 定理 若 X 是 $(\Omega, \mathcal{F}, \mathbf{P})$ 到 $(E, \mathcal{B})$ 上的可测映射, X 在 $\mathcal{F}$ 的子 σ 代数 $\mathcal{C}$ 下的条件概率分布 $\mathbf{P}_X^{\mathcal{C}}(\omega, B)$, $(\omega, B) \in \Omega \times \mathcal{B}$ 存在, 则对任何使 $\mathbf{E}f(X)$ 存在的 $(E, \mathcal{B})$ 上的可测函数 f, 均有

$$\mathbf{E}(f(X)|\mathcal{C})(\cdot) = \int_E f(x)\mathbf{P}_X^{\mathcal{C}}(\cdot, dx), \quad \mathbf{P}_{\mathcal{C}} - \text{a.e.}.$$

6. 定理 若 X 是由概率空间 $(\Omega, \mathcal{F}, \mathbf{P})$ 到完备可分距离可测空间 $(E, \mathcal{B})$ 的可测映射, 则对任何 $\mathcal{F}$ 的子 σ–域 $\mathcal{C}$, 存在 X 在 $\mathcal{C}$ 之下的条件概率分布. 特别若 $X = (X_1, \cdots, X_n, \cdots)$ 为 $(\Omega, \mathcal{F}, \mathbf{P})$ 上的可数维实随机变量, $\mathcal{C}$ 为 $\mathcal{F}$ 的任一子 σ–域, 则存在 X 在 $\mathcal{C}$ 之下的条件概率分布.

为证明定理 6, 我们首先证明下述一系列引理.

7. 引理 若 $(E,\mathcal{B})$ 为距离可测空间, $\mathbf{P}$ 为其上概率测度, 则对任何 $B\in\mathcal{B}$ 及 $\varepsilon>0$ 存在开集 G 和闭集 F 使得 $F\subset B\subset G$ 且 $\mathbf{P}(G\setminus F)<\varepsilon$, 特别当 (E,ρ) 完备可分时, F 可选为紧集.

证明 若 B 为闭集, 取开集 $G_n=\{x\in E:\rho(x,B)<\frac{1}{n}\}$, 则显然有 $B=\cap_{n=1}^{\infty}G_n$, 再由概率的上连续性即得所需结论. 令

$$\Lambda=\{B\in\mathcal{B}:\forall\varepsilon>0,\exists\text{ 开集}G,\text{闭集}F,$$

$$\text{使得}F\subset B\subset G\text{ 且}\mathbf{P}(G\setminus F)<\varepsilon\}.$$

往证 Λ 为 σ 代数. 显然 Λ 对求余运算封闭. 若 $\{B_n\}\subset\Lambda$, 则任给 $\varepsilon>0$, 对一切 $n\in\mathbb{N}$, 存在开集 G_n 及闭集 F_n 使得 $F_n\subset B_n\subset G_n$ 且 $\mathbf{P}(G_n\setminus F_n)<\frac{\varepsilon}{2^{n+1}}$. 取 $G=\cup_nG_n,F=\cup_{1\leqslant k\leqslant n_0}F_k$, 其中 n_0 选择得使 $\mathbf{P}(\cup_nF_n\setminus F)<\frac{\varepsilon}{2}$. 于是 $F\subset\cup_nB_n\subset G$ 且 $\mathbf{P}(G\setminus F)<\varepsilon$. 这就证明了 Λ 对可数并封闭, 因而是 σ–代数. 故定理的第一部分获证.

若 E 完备可分, 先对 E 证明结论成立. 设 $\{x_k,k\in\mathbb{N}\}$ 为 E 的可数稠集, 则对任何 n, $E=\cup_{k=1}^{\infty}B(x_k,\frac{1}{n})$, 其中 $B(x_k,\frac{1}{n})$ 表示以 x_k 为心, $\frac{1}{n}$ 为半径的球. 对一切 $\varepsilon>0$ 存在 k_n 使 $\mathbf{P}(\cup_{k=1}^{k_n}B(x_k,\frac{1}{n})>1-\frac{\varepsilon}{2^n}$. 取 $K=\cap_n\cup_{k=1}^{k_n}\bar{B}(x_k,\frac{1}{n})$, 其中 $\bar{B}(x_k,\frac{1}{n})$ 是 $B(x_k,\frac{1}{n})$ 的闭包. 则 K 是完全有界集, 因而是紧集, 且

$$\begin{aligned}\mathbf{P}(K^c)&=\mathbf{P}(\cup_n(\cup_{k=1}^{k_n}\bar{B}(x_k,\frac{1}{n}))^c)\\&\leqslant\textstyle\sum_n\mathbf{P}((\cup_{k=1}^{k_n}\bar{B}(x_k,\frac{1}{n}))^c)\\&\leqslant\mathbf{P}((\cup_{k=1}^{k_n}B(x_k,\frac{1}{n}))^c)<\varepsilon.\end{aligned}$$

再证对任何 $B\in\mathcal{B}$ 结论成立. 由第一段的证明知, 任给 $\varepsilon>0$, 存在开集 G 和闭集 F 使得 $F\subset B\subset G$ 且 $\mathbf{P}(G\setminus F)<\frac{\varepsilon}{2}$, 又存在紧集 K 使得 $\mathbf{P}(K^c)<\frac{\varepsilon}{2}$, 于是 $K\cap F$ 是紧集且

$$\mathbf{P}(G\setminus(K\cap F))=\mathbf{P}(G\setminus F)+\mathbf{P}(F\setminus(K\cap F))<\varepsilon.\qquad\square$$

8. 定理 (E,ρ) 为可分距离空间， $\mathcal{B}$ 是由开集生成的 σ 域，则存在只有可数个元的集代数 $\mathcal{A}$ 使得 $\mathcal{B}=\sigma(\mathcal{A})$，称 $\mathcal{A}$ 为 $\mathcal{B}$ 的可数生成域.

证明 设 $\{x_k\ k\in\mathbb{N}\}$ 是 $(E,\ \rho)$ 的可数稠集，则取 $\mathcal{A}$ 为 $\{B(x_k,\frac{1}{n}):k,n\in\mathbb{N}\}$ 上的最小集代数即可. （注意：可数个集合生成的最小集代数只有可数个集合.）

9. 引理 设 $(E,\mathcal{B})$ 为距离可测空间， $\mathcal{A},\widetilde{\mathcal{A}}$ 均为 E 的集代数，且 $\mathcal{A}\subset\widetilde{\mathcal{A}}\subset\mathcal{B}$, $\sigma(\mathcal{A})=\mathcal{B}$. 若 μ 为 $\widetilde{\mathcal{A}}$ 上取有限值的有限可加测度，且对任何 $A\in\mathcal{A}$, 存在紧集序列 $\{C_n\}\subset\widetilde{\mathcal{A}},C_n\subset A$, $n\in\mathbb{N}$, 使得当 $n\to\infty$ 时 $\mu(C_n)\to\mu(A)$, 则 μ 在 $\mathcal{A}$ 上 σ 可加，因而可唯一扩张为 $\mathcal{B}$ 上的测度.

证明 任取 $\mathcal{A}$ 中两两不交的集序列 $\{A_n\}$, 且 $\cup_{n=1}^{\infty}A_n=A\in\mathcal{A}$. 对任何 $\varepsilon>0$, 存在紧集 $C\in\widetilde{\mathcal{A}}$,$C\subset A$ 使

$$\mu(\cup_{n=1}^{\infty}A_n)<\mu(C)+\varepsilon;$$

而对每一 $A_n^c(\in\mathcal{A}!)$ 存在紧集 $C_n(\in\widetilde{\mathcal{A}}!)\subset A_n^c$, 使

$$\mu(A_n^c)<\mu(C_n)+\frac{\varepsilon}{2^n}.$$

C_n^c 为开集且

$$\mu(C_n^c)<\mu(A_n)+\frac{\varepsilon}{2^n}.$$

显然 $\cup_{n=1}^{\infty}C_n^c\supset C$. 由有限覆盖定理，存在自然数 m, 使 $\cup_{n=1}^{m}C_n^c\supset C$. 于是

$$\begin{aligned}\textstyle\sum_{n=1}^{\infty}\mu(A_n)+\varepsilon&\geqslant\textstyle\sum_{n=1}^{m}\big(\mu(A_n)+\frac{\varepsilon}{2^n}\big)\\&>\textstyle\sum_{n=1}^{m}\mu(C_n^c)\geqslant\mu(\cup_{n=1}^{m}C_n^c)\\&\geqslant\mu(C)>\mu(\cup_nA_n)-\varepsilon.\end{aligned}$$

令 $\varepsilon\to0$, 即得

$$\textstyle\sum_{n=1}^{\infty}\mu(A_n)\geqslant\mu(\cup_{n=1}^{\infty}A_n).$$

相反的不等式

$$\sum_{n=1}^{\infty}\mu(A_n)\geqslant\mu(\cup_n A_n)$$

对任何有限可加测度成立. 这就证明了 μ 是 $\mathcal{A}$ 上的测度. 由测度扩张定理, 可唯一扩张为 $\mathcal{B}$ 上的测度.

10. 引理 若 $(\Omega,\mathcal{C},\mu)$ 为一测度空间, $(E,\mathcal{B})$ 为完备可分距离空间. 设 $\mathcal{B}$ 的可数生成域为 $\mathcal{A}$, $R(\omega,A),\omega\in\Omega,A\in\mathcal{B}$ 为有限非负实函数, 满足

1) 对一切 $A\in\mathcal{B}$, $R(\cdot,A)$ 为 $\mathcal{C}$–可测函数;

2) 对一切 $A\in\mathcal{A}$, 存在紧集序列 $\{C_n^A:\ n\in\mathbb{N}\}$, 每一 $C_n^A\subset A$,

$$R(\omega,A)=\lim_{n\to\infty}R(\omega,C_n^A),\quad \mu-\text{a.e.};$$

3) 对一切 $\{A_n\}\subset\mathcal{B},A_n\cap A_m=\varnothing,m\neq n$, 有

$$R(\cdot,\cup_{n=1}^{\infty}A_n)=\sum_{n=1}^{\infty}R(\cdot,A_n),\quad \mu-\text{a.e.},$$

则存在函数 $U(\omega,A),\omega\in\Omega,A\in\mathcal{B}$, 满足

4) 对一切 $\omega\in\Omega$, $U(\omega,\cdot)$ 是 $(E,\mathcal{B})$ 上的有限测度;

5) 对一切 $A\in\mathcal{B}$, $U(\cdot,A)$ 是 $\mathcal{C}$ 可测函数且

$$U(\cdot,A)=R(\cdot,A),\ \mu-\text{a.e.}.$$

证明 对每一 $A\in\mathcal{A}$, 按条件 2) 可以决定一可数紧集序列 $\{C_n\}$, 把它们都添入 $\mathcal{A}$, 再生成一个可数域 $\widetilde{\mathcal{A}}$, 往证存在一个 $\mathcal{C}$ 可测的 μ–零集 N, 使得

$$R(\omega,\cdot)\text{为}\widetilde{\mathcal{A}}\text{上的有限可加测度},\ \forall\,\omega\in N^c; \tag{1}$$

$$R(\omega,A)=\lim_{n\to\infty}R(\omega,C_n^A),\ \forall\,A\in\mathcal{A},\omega\in N^c. \tag{2}$$

事实上, 由 R 有限及 3) 知, 存在 μ 零集 (若 $M\in\mathcal{C}$ 且 $\mu(M)=0$, 称 M 为 μ 零集) N_0 使

$$R(\omega,\varnothing)=0,\ \omega\in N_0^c;$$

由于 $\widetilde{\mathcal{A}}$ 只有可数个元，不妨记

$$\widetilde{\mathcal{A}}=\{B_r:\ r\in Q\},$$

其中 Q 为有理数集. 若 $\{B_{r_1},\cdots,B_{r_m}\}\subset\widetilde{\mathcal{A}}$ 且两两不交，由 3) 知，存在 μ 零集 $N_{r_1,\cdots,r_m}$ 使

$$R(\omega,\cup_{k=1}^{m}B_{r_k})=\sum_{k=1}^{m}R(\omega,B_{r_k}),\ \omega\in N^c_{r_1,\cdots,r_m};$$

再由 2) 知，对每一 $A\in\mathcal{A}$，存在 μ 零集 N_A，使

$$R(\omega,A)=\lim_{n\to\infty}R(\omega,C_n^A),\ \omega\in N_A^c.$$

令

$$N=N_0\bigcup(\bigcup_{m=1}^{\infty}\bigcup_{\{r_1,\cdots,r_m\}\subset Q}N^c_{r_1,\cdots,r_m})\bigcup(\bigcup_{A\in\mathcal{A}}N_A).$$

则 (1),(2) 成立. 因而对一切 $\omega\in N^c$, $R(\omega,\cdot)$ 满足引理 9 的条件. 因而可唯一扩张为 $\mathcal{B}$ 上的有限测度，记作 $R_1(\omega,\cdot)$.

任取定 $\omega_0\in N^c$，令

$$U(\omega,A)=\begin{cases}R_1(\omega,A), & \omega\in N^c,\ A\in\mathcal{B};\\ R_1(\omega_0,A), & \omega\in N,\ A\in\mathcal{B}.\end{cases}$$

显然 U 满足 4), 往证 U 满足 5). 由于若以 R_1 代 U 使 5) 成立，则 U 也使 5) 成立，令

$$\Lambda=\{A\in\mathcal{B}:\ R_1(\cdot,A)\ \mathcal{C}-\text{可测且}\ R_1(\cdot,A)=R(\cdot,A),\ \mu-\text{a.e.}\}.$$

由 R 满足 3) 易证 Λ 为包含 $\mathcal{A}$ 的集代数且为 $\lambda-$系. 因而 $\Lambda=\mathcal{B}$，即 5) 成立. □

引理 10 不仅对证明定理 6 有用，而且有其独特作用.

11. 定理 6 的证明.

证明 对每一 $B\in\mathcal{B}$, 对事件 $\{X\in B\}$ 取定一个条件期望的版本

$$\mathbf{P}(X\in B|\mathcal{C})(\omega),\ \omega\in\Omega.$$

取 $(\Omega, \mathcal{C}, \mathbf{P}_{\mathcal{C}})$ 为引理 10 中的 $(\Omega, \mathcal{C}, \mu)$ 令

$$R(\omega, B) := \mathbf{P}(X \in B|\mathcal{C})(\omega),\ (\omega, B) \in \Omega \times \mathcal{B}.$$

由条件期望的性质立知 R 满足引理 10 的条件 1) 和 3). 由引理 7 及 (E, ρ) 的完备可分性知对 $\mathcal{B}$ 的可数生成域中的任一元 A, 存在紧集序列 $\{C_n : n \in \mathbb{N}\}$, 每一 $C_n \subset A$, 使 $\lim_{n\to\infty} \mathbf{P}(A \setminus C_n) = 0$. 即 $I_{C_n} \uparrow I_A$, $\mathbf{P}$ – a.e., 由条件概率的单调收敛性知存在 $\mathbf{P}_{\mathcal{C}}$–零集 N, 当 $\omega \notin N$ 时, $R(\omega, A) = \lim_{n\to\infty} R(\omega, C_n)$; 因而引理 10 的条件 2) 也成立. 于是存在满足引理 10 条件 4), 5) 的函数 U.

$$\mathbf{P}_X^{\mathcal{C}}(\omega, B) := U(\omega, B), \quad , \forall\, (\omega, B) \in \Omega \times \mathcal{B},$$

就是 X 在 $\mathcal{C}$ 之下的条件概率分布.

特别当 X 为无穷随机变量序列时, $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}})$ 可以看作由距离

$$d(x, y) := \sum_{n=1}^{\infty} \frac{1}{2^n} |x_n - y_n|,\ x, y \in \mathbb{R}^{\mathbb{N}}$$

决定的 Borel 可测空间, 由习题 2.4.10 知它是完备可分距离空间, 因而 X 在 $\mathcal{C}$ 之下的条件概率分布存在. □

12. 定理 若 $(E, \mathcal{B})$ 为完备可分距离空间, $\mathbf{P}$ 为其上概率测度, 则对任何子 σ 域 $\mathcal{C} \subset \mathcal{B}$, $\mathcal{B}$ 在 $\mathcal{C}$ 之下的正则条件概率总存在.

证明 在定理 6 的证明中取 $R(x, B) := \mathbf{P}(B|\mathcal{C})(x)$ 即可.

13. 定理 若 $(E_k, \mathcal{B}_k), k = 1, \cdots, n$ 是完备可分距离可测空间, $(E^k, \mathcal{B}^k)$, 是前 k 个可测空间的乘积可测空间, $k = 2, \cdots, n$. 则对其上任意概率测度 $\mathbf{P}$, 存在 $(E_1, \mathcal{B}_1)$ 上概率测度 P_1 及 $E^{k-1} \times \mathcal{B}_k$ 上转移概率 $P_k(x_1, \cdots, x_{k-1}, B), (x_1, \cdots, x_{n-1}) \in E^{k-1}$, $B \in$

$\mathcal{B}_k\ ,k=2,\cdots,n$, 使

$$\mathbf{P}(B)=\int_E\int_E\cdots\int_E I_B(x_1,\cdots,x_n) \qquad B\in\mathcal{B}^n \tag{1}$$
$$P_n(x_1,\cdots,x_{n-1},dx_n)\cdots P_2(x_1,dx_2)P_1(dx_1).$$

证明 首先, (1) 成立, 等价于

$$\int_{E^n} f(x_1,\cdots,x_n)\,d\mathbf{P} \tag{2}$$
$$=\int_{E_1}\int_{E_2}\cdots\int_{E_n} f(x_1,\cdots,x_n) \qquad \forall\, 0\leqslant f\in\mathcal{B}^n$$
$$P_n(x_1,\cdots,x_{n-1},dx_n)\cdots P_2(x_1,dx_2)P_1(dx_1).$$

令

$$\mathcal{C}_{n-1}:=\{B^{n-1}\times E_n:\ B^{n-1}\in\mathcal{B}^{n-1}\},$$

则由定理 12 知有一在 $\mathcal{C}_{n-1}$ 之下的正则条件概率

$$\mathbf{P}^{\mathcal{C}_{n-1}}(x,A),\quad (x,A)\in E^n\times\mathcal{B}^n$$

存在. 由于对任意 $(x_1,\cdots,x_{n-1}),(x_1,\cdots,x_{n-1})\times E_n$ 是 $\mathcal{C}_{n-1}$ 的原子, 故由定理 3.5 知等式

$$P_n(x_1,\cdots,x_{n-1},B_n):=\mathbf{P}^{\mathcal{C}_{n-1}}(x_1,\cdots,x_n,E^{n-1}\times B_n)$$

定义了 $E^{n-1}\times\mathcal{B}_n$ 上的函数, 而且由正则条件概率的定义易知它是 $(E^{n-1},\mathcal{B}^{n-1})$ 到 $(E_n,\mathcal{B}_n)$ 上的转移概率.

我们把 $\mathbf{P}$ 在 $\mathcal{C}_{n-1}$ 上的限制看作 $(E^{n-1},\mathcal{B}^{n-1})$ 上的概率, 记作 $\mathbf{P}_{1,\cdots,n-1}$, 由于对任何 $B_k\in\mathcal{B}_k,k=1,\cdots,n$,

$$\int_{\prod_{k=1}^{n-1}B_k} P_n(x_1,\cdots,x_{n-1},B_n)\,d\mathbf{P}_{1,\cdots,n-1}$$
$$=\int_{\prod_{k=1}^{n-1}B_k\times E_n}\mathbf{P}^{\mathcal{C}_{n-1}}(x_1,\cdots,x_n,E^{n-1}\times B_n)\,d\mathbf{P} \tag{3}$$
$$=\mathbf{P}\big((\textstyle\prod_{k=1}^{n-1}B_k\times E_n)\cap(E^{n-1}\times B_n)\big)$$
$$=\mathbf{P}(B_1\times\cdots\times B_n).$$

由测度扩张定理知 $\mathbf{P}$ 由 P_n 及 $\mathbf{P}_{1,\cdots,n-1}$ 唯一决定. 这就证明了当 $n=2$ 时 (1) 成立. 对 $\mathbf{P}_{1,\cdots,n-1}$ 重复上面的讨论并继续做下去即知 $\mathcal{B}^n$ 上的概率 $\mathbf{P}$ 由转移概率序列

$$P_1(B_1), P_2(x_1, B_2), \cdots, P_n(x_1, \cdots, x_{n-1}, B_n)$$
$$B_k \in \mathcal{B}_k, x_k \in E_k, k=1, \cdots, n$$

决定. 其中

$$P_k(x_1, \cdots, x_{k-1}, B_k) = P_{1,\cdots,k}^{\mathcal{C}_{k-1}}(x_1, \cdots, x_k, E^{k-1} \times B_k),$$
$$k=2, \cdots, n.$$

是 $E_{k-1} \times \mathcal{B}_k$ 在 $\mathcal{C}_{k-1} = \{B^{k-1} \times E_k : B^{k-1} \in \mathcal{B}^{k-1}\}$ 之下关于 $\mathbf{P}_{1,\cdots,k}$ 的正则条件概率. 而

$$\mathbf{P}_{1,\cdots,n}(B^n) = \mathbf{P}(B^n),\ B^n \in \mathcal{B}^n;$$
$$\mathbf{P}_{1,\cdots,k}(B^k) = \mathbf{P}_{1,\cdots,k+1}(B^k \times E_{k+1}),\ B^k \in \mathcal{B}^k;$$
$$k=1, \cdots, n-1.$$

以下使用归纳法证明 (1) 式成立. 由 (3) 及测度扩张定理知当 $n=2$ 时 (1) 成立. 若 (1) (等价地 (2)) 对 $n=k-1$ 成立, 往证对 k 成立. 设 $B^k = B^{k-1} \times B_k$, 同证明 (3) 一样的方法可知对任何 $B^{k-1} \in \mathcal{B}^{k-1}$, $B_k \in \mathcal{B}_k$,

$$\mathbf{P}_{1,\cdots,k}(B^{k-1} \times B_k) = \int_{B^{k-1}} P_k(x_1, \cdots, x_{k-1}, B_k)\, d\mathbf{P}_{1,\cdots,k-1}$$
$$= \int_{E^{k-1}} I_{B^{k-1}}(x_1, \cdots, x_{k-1}) P_k(x_1, \cdots, x_{k-1}, B_k)\, d\mathbf{P}_{1,\cdots,k-1},$$

其中 $I_{B^{k-1}} P_k(x_1, \cdots, x_{k-1}, B_k)$, $(x_1, \cdots, x_{k-1}) \in E^{k-1}$ 是 $\mathcal{B}^{k-1}$

非负可测函数. 故由 (2) 对 $k-1$ 成立知

$$\begin{aligned}\mathbf{P}_{1,\cdots,k}(B^{k-1}\times B_k) &= \int_{E_1}\int_{E_2}\cdots\int_{E_{k-1}} I_{B^{k-1}}(x_1,\cdots,x_{k-1})\\ &\quad\times P_k(x_1,\cdots,x_{k-1},B_k)P_{k-1}(x_1,\cdots,x_{k-2},dx_{k-1})\\ &\quad\quad\cdots P_2(x_1,dx_2)P_1(dx_1)\\ &= \int_{E_1}\int_{E_2}\cdots\int_{E_{k-1}}\int_{E_k} I_{B^{k-1}\times B_k}(x_1,\cdots,x_k)\\ &\quad\times P_k(x_1,\cdots,x_{k-1},dx_k)P_{k-1}(x_1,\cdots,x_{k-2},dx_{k-1})\\ &\quad\quad\cdots P_2(x_1,dx_2)P_1(dx_1).\end{aligned}$$

因此 (1) 式对 $n=k$ 且 $B^k=B^{k-1}\times B_k$, $B^{k-1}\in\mathcal{B}^{k-1}$, $B_k\in\mathcal{B}_k$ 成立. 由测度扩张定理及 (1) 式右端是一个测度 (为什么?) 知 (1) 对 $n=k$ 成立. 因此 (1) 对一切自然数 n 成立. □

下面介绍在随机过程论中有基础意义的 Kolmogorov 相容性定理. 首先引入概率分布族相容的概念.

14. 定义 设 T 为一无穷集, 对每一 $t\in T$, 有一可测空间 $(\Omega_t,\mathcal{F}_t)$, 若对 T 的某一有限子集类中的每一 T_N, 有一概率场 $(\Omega^{T_N},\mathcal{F}^{T_N},\mathbf{P}^{T_N})$, 其中 $\Omega^{T_N}=\prod_{t\in T_N}\Omega_t$, $\mathcal{F}^{T_N}=\prod_{t\in T_N}\mathcal{F}_t$, 且它们具有下列性质: 对于 T 的这个有限子集中的任何两个元 T_N, T_N', $T_N\subset T_N'$ 及任一 $A^{T_N}\in\mathcal{F}^{T_N}$,

$$\mathbf{P}^{T_N}(A^{T_N})=\mathbf{P}^{T_N'}(A^{T_N}\times\textstyle\prod_{t\in T_N'-T_N}\Omega_t),$$

则称 **诸概率场是相容的.**

15. 定理 (Колмогоров 相容性定理) 设 $(E,\mathcal{B})$ 为完备可分距离可测空间, T 为一无穷集, 对 T 的任一有限子集 T_N, 有一概率场 $(E^{T_N},\mathcal{B}^{T_N},\mathbf{P}^{T_N})$, 它们是相容的. 则在 $(E^T,\mathcal{B}^T)$ 上存在唯一概率测度 $\mathbf{P}$ 使

$$\begin{aligned}&\mathbf{P}(A^{T_N}\times E^{T\setminus T_N})=\mathbf{P}^{T_N}(A^{T_N}),\\ &\forall\, A^{T_N}\in\mathcal{B}^{T_N},\quad \text{有限集}\, T_N\subset T.\end{aligned}\tag{1}$$

证明 (i) 当 T 是可数无穷集时，不妨设 $T=\mathbb{N}$，于是概率场 $(E^n,\mathcal{B}^n,\mathbf{P}^n), n\in\mathbb{N}$ 是相容的，其中 $E^n=\prod_{i=1}^n E_i$，$\mathcal{B}^n=\prod_{i=1}^n \mathcal{B}_i$，$\mathbf{P}^n=\mathbf{P}^{\{1,\cdots,n\}}$. 因此由定理 **13** 的证明知它们决定一个转移概率序列

$$P_1(B_1), P(x_1,B_2),\cdots,P_n(x_1,\cdots,x_{n-1},B_n),\cdots$$
$$x_k\in E,\ B_k\in\mathcal{B},\ k=1,\cdots,n,$$

其中

$$P_n(x_1,\cdots,x_{n-1},B_n)=(\mathbf{P}^n)^{\mathcal{C}_{n-1}}(x_1,\cdots,x_n,E^{n-1}\times B_n)$$
$$n=2,\cdots.$$

而 $(\mathbf{P}^n)^{\mathcal{C}_{n-1}}$ 是在 $\mathcal{C}_{n-1}$ 之下关于 $\mathbf{P}^n$ 的正则条件概率. 于是由 **Tulcea** 定理知在 $(E^{\mathbb{N}},\mathcal{B}^{\mathbb{N}})$ 上有唯一确定的概率 $\mathbf{P}^{\mathbb{N}}$ 存在，使得

$$\mathbf{P}^{\mathbb{N}}(B^n\times\prod_{k=1}^{\infty}E_k)=\int_{E_1}\int_{E_2}\cdots\int_{E_n}I_{B^n}(x_1,\cdots,x_n)$$
$$\times P_n(x_1,\cdots,x_{n-1},dx_n)\cdots P_2(x_1,dx_2)\mathbf{P}^1(dx_1),$$
$$\forall\, B^n\in\mathcal{B}^n.$$

再由定理 **13** 知上式右端即 $\mathbf{P}^n(B^n)$.

对这种情形还需证明对任意有限集 $T_N\subset\mathbb{N}$ 有

$$\mathbf{P}^{\mathbb{N}}\left(B^{T_N}\times\prod_{n\in\mathbb{N}\setminus T_N}E_n\right)=\mathbf{P}^{T_N}(B^{T_N}),\ B^{T_N}\in\mathcal{B}^{T_N}.$$

事实上，存在 m 使 $T_N\subset\{1,\cdots,m\}$，于是

$$B^{T_N}\times\prod_{n\in\mathbb{N}\setminus T_N}E_n=\left(B^{T_N}\prod_{k\in\{1,\cdots,m\}\setminus T_N}E_k\right)\times\prod_{n=m+1}^{\infty}E_n.$$

因而由 $\mathbf{P}^{\mathbb{N}}$ 的定义及 $\mathbf{P}^{T_N}, T_N\subset\mathbb{N}$ 的相容性即得

$$\begin{aligned}&\mathbf{P}^{\mathbb{N}}\left(B^{T_N}\times\prod_{n\in\mathbb{N}\setminus T_N}E_n\right)\\&=\mathbf{P}^m\left(B^{T_N}\times\prod n\in\{1,\cdots,n\}\setminus T_N E_n\right)\\&=\mathbf{P}^{T_N}(B^{T_N}).\end{aligned}$$

(ii) 设 T 是不可数集，任取 T 的一个可数无穷子集 T_c，则由 **(i)** 知在 $\mathcal{B}^{T_c}$ 上有唯一的概率 $\mathbf{P}^{T_c}$ 存在，使得对任何以 $B^{T_N} \in \mathcal{B}^{T_N}$,($T_N \subset T_c$) 为底的柱集 $B^{T_N} \times \prod_{t\in T_c\setminus T_N} E_t$，

$$\mathbf{P}^{T_c}(B^{T_N} \times \prod_{t\in T_c\setminus T_N} E_t) = \mathbf{P}^{T_N}(B^{T_N}).$$

由引理 **6.3.3** 知对任一 $B^T \in \mathcal{B}^T$，必有一 T 的可数子集 T_c，使得 $B^T = B^{T_c} \times \prod_{t\in T\setminus T_c} E_t$，令

$$\mathbf{P}^T(B^T) := \mathbf{P}^{T_c}(B^{T_c}).$$

几乎逐字重复无穷乘积概率定理证明的有关部分，可知上式定义的 $\mathbf{P}^T$ 是 $(E^T, \mathcal{B}^T)$ 上唯一满足 **(1)** 的概率.

习题

1. 设 X_T 是 $(\Omega, \mathcal{F}, \mathbf{P})$ 到 $(\mathbb{R}^T, \mathcal{B}^T)$ 上的可测映射 $\mathcal{C}$ 是 $\mathcal{F}$ 的子 σ 代数，则 $\mathbf{P}^{\mathcal{C}}(\omega, B)$, $(\omega, B) \in \Omega \times \mathcal{B}^T$ 是 X_T 在 $\mathcal{C}$ 之下的条件概率分布的充分必要条件是下述二条件同时成立:

1) $\mathbf{P}^{\mathcal{C}}(\omega, B)$ 是 $(\Omega, \mathcal{C})$ 到 $(\mathbb{R}^T, \mathcal{B}^T)$ 的转移概率；

2) 对一切 $B^T \in \mathcal{B}^T$, $C \in \mathcal{C}$,

$$\int_C \mathbf{P}^{\mathcal{C}}(\cdot, B^T) d\mathbf{P} = \mathbf{P}(C \cap \{X_T \in B^T\}).$$

2. 设 $X = (X_1, \cdots, X_n)$ 是 $(\mathbb{R}^n, \mathcal{B}^n, \mathbf{P})$ 上的样本函数，具有 n 维正态分布 $N(0, D)$，令

$$Y(x) = xD^{-1}x', \quad x \in \mathbb{R}^n,$$

试求 X 在 $\sigma(Y)$ 之下的条件概率分布.

3. 称随机变量序列 $X_n, n \in \mathbb{N}$ 为一马氏序列，如果对任意 $B \in \mathcal{B}$, $(x_1, \cdots, x_{n-1}) \in \mathbb{R}^{n-1}$, $n = 2, 3, \cdots$，有

$$\begin{aligned}&\mathbf{P}(X_n \in B | X_1 = x_1, \cdots, X_{n-1} = x_{n-1}) \\ &= \mathbf{P}(X_n \in B | X_{n-1} = x_{n-1}), \text{a.e.}.\end{aligned}$$

若给定一组转移概率序列

$$P_1(B_1),\ B_1 \in \mathcal{B},\ P_n(x_{n-1}, B_n),$$
$$B_n \in \mathcal{B},\ x_{n-1} \in \mathbb{R},\ n = 2, 3, \cdots,$$

则有一定义在概率空间 $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}}, \mathbf{P}^{\mathbb{N}})$ 上的马氏序列 X_n, $n \in \mathbb{N}$, 使得 X_n 在 X_{n-1} 之下的条件概率

$$\mathbf{P}^{\mathbb{N}}(X_n \in B | X_{n-1} = x_{n-1}) = P_n(x_{n-1}, B), \mathbf{a.e.},$$
$$B \in \mathcal{B}, \quad x_{n-1} \in \mathbb{R}, \quad n = 2, 3, \cdots;$$

而

$$\mathbf{P}^{\mathbb{N}}(X_1 \in B) = P_1(B), B \in \mathcal{B}.$$

(提示：将此处的 $P_n(x_{n-1}, B_n)$ 看成是定理 **15** 证明中的 $P_n(x_1, \cdots, x_{n-1}, B_n)$.)

第八章 收敛概念

在测度论和概率论的研究中，一个重要的内容是各种收敛性. 本章的目的就是介绍测度空间 $(\Omega,\mathcal{F},\mu)$ 到距离可测空间 $(E,\mathcal{B})$ ((E,d) 是距离空间，$\mathcal{B}$ 是由 E 中开集生成的 σ–域) 上函数 (即映射) 列的各种收敛概念它们之间的关系以及测度的各种收敛性.

§1. 几乎处处收敛

由数列极限概念直接引出的函数序列收敛的概念自然是逐点收敛的概念，即设 $f, f_n: \Omega\mapsto E$, 对一切 $\omega\in\Omega$, 当 $n\to\infty$ 时，$f_n(\omega)\to f(\omega)$, 将此收敛概念减弱一点，自然就给出

1. 定义 设 $(\Omega,\mathcal{F},\mu)$ 是一测度空间，(E,d) 是一距离空间，$f_n: \Omega\mapsto E, n\in\mathbb{N}$, 是它上面定义的可测函数列 (更广一点可以是 μ–a.e. 可测函数，此处及本节以下所说可测函数都是指 $f:\Omega\mapsto E$, 对一切 $B\in\mathcal{B}, \{f\in B\}\in\mathcal{F}$). 若存在 $f:\Omega\mapsto E$, 及一 μ–零集 N, 使

$$f_n(\omega)\to f(\omega),\qquad \forall\omega\in N^c. \tag{1}$$

(即 $\mu\big(\{\omega:\ \overline{\lim}_{n\to\infty} d(f_n(\omega),f(\omega))\neq 0\}\big)=0$), 则称 $\{f_n:\ n\in\mathbb{N}\}\mu$–a.e. 收敛于 f, 简记作 $f_n\xrightarrow{\text{a.e.}}f$(或 $f_n\to f$, μ–a.e.). 若有一 μ–零集 N 使对一切 $\varepsilon>0$, 对一切 $\omega\in N^c$, 存在 $n_0=n_0(\omega)$ 使当 $n,m\geq n_0$ 时，$d(f_n(\omega),f_m(\omega))<\varepsilon$ (即 $\mu\big(\{\omega:\ f_n(\omega), n\in\mathbb{N}$, 不是 E 中基本列 $\}\big)=0$), 则称 $\{f_n:\ n\in\mathbb{N}\}$ a.e. 相互收敛，简记作 $d(f_n,f_m)\to 0\ (n,m\to\infty)$, μ–a.e..

由定义容易看出 (留给读者作为习题):

1) 若 $f_n \xrightarrow{\text{a.e.}} f$, $\{f_{n_k}, k \in \mathbb{N}\}$ 为 $\{f_n: n \in \mathbb{N}\}$ 的任一子序列, 则 $f_{n_k} \xrightarrow{\text{a.e.}} f$, 当 $k \to \infty$ 时,

2) 若 $f_n \xrightarrow{\text{a.e.}} f$, 且 $f_n \xrightarrow{\text{a.e.}} f'$, 则 $f = f'$ $\mu - a.e.$, 这就是说几乎处处收敛的函数列的极限函数是 a.e. 唯一决定的.

但是读者应该注意, 在数学分析中有下列的命题: 函数族 $\{f_t: t \in (0,1)\}$ 当 $t \to 0$ 时收敛的充要条件是它的任意可数子列 $\{f_{t_n}: t_n \to 0\}$ 收敛, 对于 a.e. 收敛来说, 此命题的必要条件仍然成立, 但充分条件却不再成立, 因为每一可数子列可能有一不收敛的例外 μ–零集, 这些 μ–零集可以不同, $f_t(\omega)$, $t \to 0$, 不收敛的例外集是这些例外集的并, 而 $t \to 0$ 的可数子列可以有不可数多, 于是 $f_t(\omega)$, $t \to 0$, 不收敛的例外集就可能不是 μ–零集了.

3) $f_n \xrightarrow{\text{a.e.}} f$, $g_n = f_n$, μ–a.e., $g = f$, μ–a.e., 则 $g_n \to g$, μ–a.e. (注意需要是同一 μ).

4) 设 $f_k^{(n)}: \Omega \mapsto \mathbb{R}$, $n \in \mathbb{N}$, $k = 1, \cdots, m$, 若 $f_k^{(n)} \to f_k$ μ–a.e., $k = 1, \cdots, m$, $G(x_1, \cdots, x_m)$ 是 $\mathbb{R}^m$ 上的连续函数, 则

$$G(f_1^{(n)}, \cdots, f_m^{(n)}) \to G(f_1, \cdots, f_m), \qquad \mu - \text{a.e.}.$$

特别若 $f_n \to f$, μ–a.e., $g_n \to g$, μ–a.e., 则 $f_n \pm g_n \to f \pm g$, μ–a.e., $f_n g_n \to fg$, μ–a.e., $cf_n \to cf$, μ–a.e. (c 为任意常数); $|f_n|^r \to |f|^r$, μ – a.e. ($r > 0$ 任意), 若还有 $\{g_n = 0\}$(当 n 充分大), $\{g = 0\}$ 都是 μ–零集, 则 $f_n/g_n \to f/g$, μ – a.e..

5) 若 (E, d) 完备, 则 f_n, μ – a.e. 收敛于 (某一) f 的充要条件是 $\{f_n\}$, μ – a.e. 相互收敛, 特别当 (E, d) 为 $\mathbb{R}^n$ 时, 上述结论成立.

2. 读者应该学会对于概率论中下列探讨问题的方式: 先在给定样本点 ω 时, 将函数列的值列 $\{f_n(\omega): n \in \mathbb{N}\}$ 当作点列来讨

论；然后将有关的论断转换成关于 ω 的命题，下面关于 a.e. 收敛的判别条件就是这种方法的一个好例证. 为此我们先讨论 $\{f_n\}$, $n \in \mathbb{N}$ 收敛集的表达.

读者一定熟悉：$f_n(\omega) \to f(\omega)$, 等价于对任何 $\varepsilon > o$, 存在 $m = m(\omega)$, 使对一切 $n \geq m$ 有 $d(f_n(\omega), f(\omega)) < \varepsilon$, 也等价于下列命题：对任意取定的 $0 < \varepsilon_k \to 0$, $k \in \mathbb{N}$ $(k \to \infty)$, 对任何 $k \in \mathbb{N}$, 存在 $m = m(k)$ 使对一切 $n \geq m$, $d(f_n(\omega), f(\omega)) < \varepsilon_k$, 由此很容易知道

$$
\begin{aligned}
\{f_n \to f\} &= \{\omega:\ f_n(\omega) \to f(\omega)\} \\
&= \bigcap_{\varepsilon>0} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \{\omega:\ d(f_n(\omega), f(\omega)) < \varepsilon\} \\
&= \bigcap_{\varepsilon>0} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \{d(f_n, f) < \varepsilon\} \\
&= \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \{\omega:\ d(f_n(\omega), f(\omega)) < \varepsilon_k\} \\
&= \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \{d(f_n, f) < \varepsilon_k\}.
\end{aligned}
\tag{1}
$$

由于 $d(x,y)$ 是 $E \times E$ 上的连续函数所以 $\{f_n \to f\} \in \mathcal{F}$.

类似地可得

$$
\begin{aligned}
\{d(f_n, f_m) \to 0\} &= \bigcap_{\varepsilon>0} \bigcup_{m=1}^{\infty} \bigcap_{\nu=1}^{\infty} \{d(f_{n+\nu}, f_n) < \varepsilon\} \\
&= \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{\nu=1}^{\infty} \{d(f_{n+\nu}, f_n) < \varepsilon_k\},
\end{aligned}
\tag{2}
$$

其中 ε_k, $k \in \mathbb{N}$ 是任意取定趋于 0 的正数列，显然 $\{d(f_n, f_m) \to 0\} \in \mathcal{F}$.

今往证明下面的 a.e. 收敛判别条件.

3. 定理 设 f_n, $n \in \mathbb{N}$, f 都是 $(\Omega, \mathcal{F}, \mu)$ 上的可测函数，则

(i) $f_n \xrightarrow{\text{a.e.}} f$, $n \to \infty$, 的充要条件是

$$\mu\Big(\bigcap_{n=1}^{\infty}\bigcup_{\nu=1}^{\infty}\big\{d(f_{n+\nu}, f) \geq \varepsilon\big\}\Big) = 0, \qquad \forall\, \varepsilon > 0. \tag{1}$$

特别当 μ 为有限测度 (自然对概率也适用) 时, $f_n \xrightarrow{\text{a.e.}} f$ 当且仅当

$$\forall\, \varepsilon > 0, \qquad \mu\Big(\bigcup_{\nu=1}^{\infty}\big\{d(f_{n+\nu}, f) \geq \varepsilon\big\}\Big) \to 0, \quad (n \to \infty). \tag{2}$$

(ii) $d(f_n, f_m) \xrightarrow{\text{a.e.}} 0$, 的充要条件是

$$\forall\, \varepsilon > 0, \qquad \mu\Big(\bigcap_{n=1}^{\infty}\bigcup_{\nu=1}^{\infty}\big\{d(f_{n+\nu}, f_n) \geq \varepsilon\big\}\Big) = 0. \tag{3}$$

特别当 μ 为有限测度时, $d(f_n, f_m) \xrightarrow{\text{a.e.}} 0$ 当且仅当

$$\forall\, \varepsilon > 0, \qquad \mu\Big(\bigcup_{\nu=1}^{\infty}\big\{d(f_{n+\nu}, f_n) \geq \varepsilon\big\}\Big) \to 0, \quad (n \to \infty). \tag{4}$$

(iii) 当 μ 为一般测度时, (2), (4) 也是 (1), (3) 成立的充分条件.

证明 (i) 若 $f_n \xrightarrow{\text{a.e.}} f$, 则对任何 $\varepsilon > 0$, 由 (2.1)

$$\begin{aligned}\mu\Big(\bigcap_{n=1}^{\infty}\bigcup_{\nu=1}^{\infty}\big\{d(f_{n+\nu}, f) \geq \varepsilon\big\}\Big) &\leq \mu\Big(\bigcup_{\varepsilon>0}\bigcap_{n=1}^{\infty}\bigcup_{\nu=1}^{\infty}\big\{d(f_{n+\nu}, f) \geq \varepsilon\big\}\Big) \\ &= \mu(\{f_n \to f\}^c) = 0.\end{aligned}$$

反之, 若 (1) 成立, 则任取 $\{\varepsilon_k\}$, $0 < \varepsilon_k \to 0$, 于是由 (2.1) 知

$$\begin{aligned}\mu(\{f_n \to f\}^c) &= \mu\Big(\bigcup_{k=1}^{\infty}\bigcap_{n=1}^{\infty}\bigcup_{\nu=1}^{\infty}\big\{d(f_{n+\nu}, f) \geq \varepsilon_k\big\}\Big) \\ &\leq \sum_{k=1}^{\infty}\mu\Big(\bigcap_{n=1}^{\infty}\bigcup_{\nu=1}^{\infty}\big\{d(f_{n+\nu}, f) \geq \varepsilon_k\big\}\Big) = 0.\end{aligned}$$

故 $f_n \xrightarrow{\text{a.e.}} f$ 当且仅当 (1) 成立.

注意 $\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu},f)\geq\varepsilon\}$, 对 n 是一不升集列, 故当 μ 为有限测度时,

$$\mu\Big(\bigcap_{n=1}^{\infty}\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu},f)\geq\varepsilon\}\Big)=\lim_{n\to\infty}\mu\Big(\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu},f)\geq\varepsilon\}\Big).$$

因而, 此时 (1) 与 (2) 等价, 由此即得 (i) 的第二个结论. 由于对一般测度 μ, 对一切 $n\in\mathbb{N}$, 有

$$\mu\Big(\bigcap_{n=1}^{\infty}\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu},f)\geq\varepsilon\}\Big)\leq\mu\Big(\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu},f)\geq\varepsilon\}\Big).$$

故对一般测度 μ, (2) 也是 $f_n\to f$, μ–a.e. 的充分条件.

(ii) $d(f_n,f_m)\xrightarrow{\text{a.e.}}0$ 的充要条件是 (3) 成立的证明与 (i) 的证明类似. 今证 (ii) 的后一部分. 当 μ 为有限测度时, (3) 等价于

$$(5)\qquad\forall\varepsilon>0,\qquad\lim_{m\to\infty}\mu\Big(\bigcap_{n=1}^{m}\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu},f_n)\geq\varepsilon\}\Big)=0.$$

对一切 $k\leqslant n$,

$$\begin{aligned}&\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu},f_n)\geqslant\varepsilon\}\\ \subset&\bigcup_{\nu=1}^{\infty}\Big(\{d(f_{n+\nu},f_k)\geqslant\frac{\varepsilon}{2}\}\bigcup\{d(f_k,f_n)\geqslant\frac{\varepsilon}{2}\}\Big)\\ \subset&\bigcup_{\nu=1}^{\infty}\{d(f_{k+\nu},f_k)\geqslant\frac{\varepsilon}{2}\}.\end{aligned}$$

所以,

$$\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu},f_n)\geqslant\varepsilon\}\subset\bigcap_{k=1}^{n}\bigcup_{\nu=1}^{\infty}\{d(f_{k+\nu},f_k)\geqslant\varepsilon\}.$$

由 (5) 即得 (4). 对一般测度 μ, (4) 蕴含 (3) 是显然的. □

注 在概率论中, 由于习惯上称 Ω 为必然事件, 因此对 r.v. $X_n\to X$, $\mathbf{P}$ – a.e. (相应地 $d(X_n,X_m)\to 0$, $\mathbf{P}$– a.e.), 很多文献上称为 "几乎必然 (almost sure) 收敛", 记作 $X_n\xrightarrow{\text{a.s.}}X$, (相应地 $d(X_n,X_m)\to 0$, a.s.).

$\{f_n,\ n\in\mathbb{N}\}$ a.e. 收敛有一个常用的充分条件，为了叙述和证明它，先来叙述一个关于事件列 A_n, $n\in\mathbb{N}$, 无穷次发生 (记作 A_ni.o.) 的充分条件，即著名的

4. 引理 (Borel-Cantelli) 设 $(\Omega,\mathcal{F},\mu)$ 是一测度空间， $\{A_n:n\in\mathbb{N}\}\subset\mathcal{F}$, 若

$$\sum_{n=1}^{\infty}\mu(A_n)<\infty, \tag{1}$$

则 (记 $\{A_n\text{ i.o.}\}=\{\omega:\ \omega$ 属于 A_n, $n\in\mathbb{N}$, 中的无穷个 $\}$)

$$\mu\Big(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k\Big)=\mu(A_n\text{ i.o.})=0. \tag{2}$$

证明 显然 $\{A_n\text{ i.o.}\}=\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k$, 于是对一切 $n\in\mathbb{N}$,

$$\mu\Big(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k\Big)\le\mu\Big(\bigcup_{k=n}^{\infty}A_k\Big)\le\sum_{k=n}^{\infty}\mu(A_k), \tag{3}$$

由 (1) 知 $\lim_{n\to\infty}\sum_{k=n}^{\infty}\mu(A_k)=0$, 故在 (3) 中令 $n\to\infty$ 即得 (2). □

由 Borel-Cantelli 引理立得

5. 引理 设 f, f_n, $n\in\mathbb{N}$, 为可测函数， $\{\varepsilon_k:\ k\in\mathbb{N}\}$ 为正数列，若 $\varepsilon_k\to0$ 且

$$\sum_{k=1}^{\infty}\mu(d(f_k,f)\ge\varepsilon_k)<\infty, \tag{1}$$

则 $f_n\xrightarrow{\text{a.e.}}f$.

证明 因为对一切 $\varepsilon>0$, 存在 $n_0=n_0(\varepsilon)$, 使对一切 $n\ge n_0$, $\varepsilon_n<\varepsilon$,

$$\begin{aligned}\mu\Big(\bigcap_{n=1}^{\infty}\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu},f)\ge\varepsilon\}\Big)&\le\sum_{k=n}^{\infty}\mu(\{d(f_k,f)\ge\varepsilon\})\\&\le\sum_{k=n}^{\infty}\mu\left(\{d(f_k,f)\ge\varepsilon_k\}\right).\end{aligned}$$

在上式中令 $n \to \infty$, 由 (1) 知 (3.1) 成立, 因而由定理 3 知 $f_n \xrightarrow{\text{a.e.}} f$. □

对相互收敛也有类似的充分条件, 即

6. 引理 设 f_n, $n \in \mathbb{N}$ 为 $(\Omega, \mathcal{F})$ 到 $(E, \mathcal{B})$ 的可测映射, δ_n, $n \in \mathbb{N}$ 为正数列, 若 $\sum_{k=1}^{\infty} \delta_k < \infty$, 且

(1) $$\sum_{n=1}^{\infty} \mu(d(f_{n+1}, f_n) \geq \delta_n) < \infty,$$

则 $\mu\big(\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu}, f_n) \geq \varepsilon\}\big) \to 0$, 对一切 $\varepsilon > 0$ 成立, 因而 f_n, $n \in \mathbb{N}$, a.e. 相互收敛.

特别, 若还有 (E, d) 完备, 则 f_n, $n \in \mathbb{N}$, 收敛.

7. 定理 (Egorov) 设 μ 为 $(\Omega, \mathcal{F})$ 上的有限测度, f, f_n, $n \in \mathbb{N}$, 为 $(\Omega, \mathcal{F})$ 到 $(E, \mathcal{B})$ 的可测映射, 若 $f_n \xrightarrow{\text{a.e.}} f$, 则对任何 $\delta > 0$, 存在 $A_\delta \in \mathcal{F}$, 使得 $\mu(A_\delta^c) < \delta$, 而

(1) $$\sup_{\omega \in A_\delta} d(f_n(\omega), f(\omega)) \to 0,$$

即所谓 f_n 几乎一致收敛于 f.

证明 f_n 在 A 上一致收敛于 f 当且仅当对任何 $m \geq 1$, 存在 $n(m)$, 使对一切 $n > n(m)$, $d(f_n(\omega), f(\omega)) < 1/m$, 对一切 $\omega \in A$ 成立, 即

$$\bigcap_{m=1}^{\infty} \bigcap_{\nu=1}^{\infty} \left\{d(f_{n(m)+\nu}, f) < \frac{1}{m}\right\} \supset A.$$

而由定理 3(ii) 知对一切 $\varepsilon > 0$,

$$\mu\Big(\bigcup_{\nu=1}^{\infty} \{d(f_{n+\nu}, f) \geq \varepsilon\}\Big) \to 0,$$

于是对一切 $m \in \mathbb{N}$, 存在 $n(m)$ 使

(2) $$\mu\Big(\bigcup_{\nu=1}^{\infty} \left\{d(f_{n(m)+\nu}, f) \geq \frac{1}{m}\right\}\Big) \leq \frac{\delta}{2^m}.$$

令

$$A_\delta^c := \bigcup_{m=1}^{\infty} \bigcup_{\nu=1}^{\infty} \left\{\omega \in \Omega:\ d(f_{n(m)+\nu}(\omega), f(\omega)) \geq \frac{1}{m}\right\}$$

于是 $\mu(A_\delta^c) < \delta$, 且在 A_δ 上 f_n 一致收敛于 f. □

习题

1. 证明引理 6.

提示: 证明对一切 $\varepsilon > 0$, 存在 n_0 使当 $n \geq n_0$ 时, $\sum_{k=n}^{\infty} \delta_k < \varepsilon$, 且对一切 $\nu \geq 1$ 有

$$\left\{d(f_{n+\nu}, f_n) \geq \varepsilon\right\} \subset \bigcup_{k=n}^{\infty} \left\{d(f_{k+1}, f_k) \geq \delta_k\right\}.$$

2. 若 $(\Omega, \mathcal{F}, \mu)$ 为测度空间, $\{B_n, n \in \mathbb{N}\} \subset \mathcal{F}$, 且存在 $\delta > 0$, 使得对每一 $n \in \mathbb{N}$, $\mu(B_n) \geqslant \delta$, 试证在 Ω 中至少存在一个点 ω 属于无穷多个 B_n.

3. 概率空间 $(\Omega, \mathcal{F}, \mathbf{P})$ 上的随机变量列 $X_n \xrightarrow{\text{a.e.}} X$ 有限, 试证对任何 $\varepsilon > 0$, 存在 $M(\varepsilon) > 0$, 使得 $\mathbf{P}(\sup|X_n| \leqslant M(\varepsilon)) \geqslant 1 - \varepsilon$.

4. 对概率空间 $(\Omega, \mathcal{F}, \mathbf{P})$ 上的任何随机变量列 $\{X_n, n \in \mathbb{N}\}$, 存在常数序列 $\{A_n, n \in \mathbb{N}\}$, 使得 $X_n/A_n \xrightarrow{\text{a.e.}} 0$.

5. 举一反例, 说明当 μ 不是有限测度时, Egorov 定理不成立.

§2　依测度收敛

比 a.e. 收敛 (或 a.s. 收敛) 较弱 (对有限测度而言) 一点的概念是 "依测度收敛", 它在概率论中有很重要的意义. 我们先给出定义.

1. 定义 设 $(\Omega, \mathcal{F}, \mu)$ 是任一给定的测度空间, (E, d) 是一距离空间, (如无特别声明, 以下所说的可测函数都是指 $(\Omega, \mathcal{F})$ 到

$(E, \mathcal{B})$ 上的可测映射，其中 $\mathcal{B}$ 为 (E, d) 的 Borel σ–代数). 设 $f, f_n, n \in \mathbb{N}$, 是可测函数. 若

$$(1) \qquad \forall \varepsilon > 0, \qquad \mu(d(f_n, f) \geq \varepsilon) \to 0 \quad (n \to \infty),$$

则称 $f_n, n \in \mathbb{N}$, **依测度 μ 收敛于** f, 记作 $f_n \xrightarrow{\mu} f \ (n \in \mathbb{N})$. 若

$$(2) \quad \forall \varepsilon > 0, \qquad \sup_{\nu} \mu\left(\left\{d(f_{n+\nu}, f_n) \geq \varepsilon\right\}\right) \to 0 \quad (n \to \infty),$$

则称 $f_n, n \in \mathbb{N}$, **依测度 μ 相互收敛**, 记作 $d(f_{n+\nu}, f_n) \xrightarrow{\mu} 0$.

注 当 $E = \mathbb{R}^m$, d 为欧氏距离时，我们按照通常习惯允许极限函数值取 $\pm\infty$, 不过由下列性质 1) 知，此时仍然不妨设极限函数取值于 E.

容易看出：

1) 设 $E = \mathbb{R}^m$, d 为欧氏距离， $f_n, n \in \mathbb{N}$, 是有限可测函数，若 $f_n \xrightarrow{\mu} f$, 则 f a.e. 有限，即有限可测函数列依测度收敛的极限是 a.e. 有限. 因而有一有限的极限函数. 即令 $\tilde{f} = f I_{\{|f|<\infty\}}$.

证明 因 f_n 有限，所以 $|f(\omega)| = \infty$, 等价于对一切 $n \geq 1$, $|f_n(\omega) - f(\omega)| = \infty$, 于是对一切 $\varepsilon > 0$,

$$\begin{aligned} \mu\left(\{|f| = +\infty\}\right) &= \mu\left(\{|f_n - f| = \infty\}\right) \\ &\leq \mu\left(\{|f_n - f| \geq \varepsilon\}\right) \to 0. \quad \square \end{aligned}$$

测度收敛也具有同普通收敛相类似的性质.

2) 若 $f_n \xrightarrow{\mu} f$, $\{f_{n_k}: k \in \mathbb{N}\}$ 是 $\{f_n: n \in \mathbb{N}\}$ 的任一子列，则 $f_{n_k} \xrightarrow{\mu} f \ (k \to \infty)$.

3) 若 $f_n \xrightarrow{\mu} f$ 且 $f_n \xrightarrow{\mu} \tilde{f}$, 则 $f = \tilde{f}$ a.e.. 即依测度收敛的极限函数是几乎唯一的.

证明 $\{f \neq \tilde{f}\} = \bigcup_{k=1}^{\infty} \{d(f, \tilde{f}) \geq \frac{1}{k}\}$. 再由 $f_n \xrightarrow{\mu} f$, 及 $f_n \xrightarrow{\mu} \tilde{f}$ 知对任何 $\varepsilon > 0$,

$$\mu\left(\left\{d(f, \tilde{f}) \geq \varepsilon\right\}\right) \leq \mu\left(\left\{d(f, f_n) \geq \frac{\varepsilon}{2}\right\}\right) + \mu\left(\left\{d(\tilde{f}, f_n) \geq \frac{\varepsilon}{2}\right\}\right).$$

令 $n \to \infty$ 即得 $\mu\left(\left\{d(f,\tilde{f}) \geq \varepsilon\right\}\right) = 0$, 故 $\mu(\{f \neq \tilde{f}\}) = 0$. □

用同样的方法可证:

4) 若 $f_n \xrightarrow{\mu} f$, $g_n = f_n$(a.e.), 则 $g_n \xrightarrow{\mu} f$.

关于依测度收敛的运算性质, 很容易证明

5) 若 $f_n : \Omega \mapsto \mathbb{R}$, $g_n : \Omega \mapsto \mathbb{R}$ 都是 $\mathcal{F}$–可测函数, 且 $f_n \xrightarrow{\mu} f$, $g_n \xrightarrow{\mu} g$, 则 $f_n \pm g_n \xrightarrow{\mu} f \pm g$.

证明 因为由 1) 知 f, g a.e. 有限, 因而可设 f, g 有限, $f \pm g$ 处处有定义. 于是对一切 $\varepsilon > 0$,

$$\begin{aligned}&\mu\left(\left\{|(f_n \pm g_n) - (f \pm g)| \geq \varepsilon\right\}\right)\\&\quad \leq \mu\left(\left\{|f_n - f| \geq \frac{\varepsilon}{2}\right\}\right) + \mu\left(\left\{|g_n - g| \geq \frac{\varepsilon}{2}\right\}\right) \to 0.\end{aligned}$$ □

测度收敛对乘除封闭的性质的证明复杂一点, 事实上可以证明下列一般的

2. 定理 设 $f_k^{(n)}$, f_k, $k = 1, \cdots, m$, $n \in \mathbb{N}$ 是 $(\Omega, \mathcal{F})$ 上的实值 (或复值) 可测函数,

$$f_k^{(n)} \xrightarrow{\mu} f_k, \qquad k = 1, 2, \cdots, m. \tag{1}$$

$g(x_1, \cdots, x_m)$ 为 n 维实空间 (或复空间) 的子集 D 上的 (实或复值) 函数, 且对一切 $\omega \in \Omega$,

$$(f_1^{(n)}(\omega), \cdots, f_m^{(n)}(\omega)) \in D, \quad (f_1(\omega), \cdots, f_m(\omega)) \in D. \tag{2}$$

1) 若 μ 是 $\mathcal{F}$ 上的测度, g 在 D 上一致连续, 则

$$g(f_1^{(n)}, \cdots, f_m^{(n)}) \xrightarrow{\mu} g(f_1, \cdots, f_m). \tag{3}$$

2) 若 μ 是 $\mathcal{F}$ 上的有限测度, D 为开集, g 在 D 上连续, 则 (3) 仍然成立.

证明 1) 较容易, 2) 复杂一点, (略, 参看 [YWL] 第二章 §5.2. 定理 3, 4).

3. 推论 设 μ 为有限测度, $f_k^{(n)} \xrightarrow{\mu} f_k$, $k = 1, 2, \cdots, m$,

$$R(x_1, \cdots, x_m) = \frac{P(x_1, \cdots, x_m)}{Q(x_1, \cdots, x_m)}, \quad P, Q \text{为多项式},$$

且对一切 $\omega \in \Omega$, $n \in \mathbb{N}$,

$$Q(f_1^{(n)}(\omega), \cdots, f_m^{(n)}(\omega)) \neq 0, \qquad Q(f_1(\omega), \cdots, f_m(\omega)) \neq 0.$$

则

$$(1) \qquad R(f_1^{(n)}, \cdots, f_m^{(n)}) \xrightarrow{\mu} R(f_1, \cdots, f_m), \qquad n \to \infty$$

特别 $f_k^{(n)} \xrightarrow{\mu} f_k$, $k = 1, 2$, $\Longrightarrow$ $f_1^{(n)} \cdot f_2^{(n)} \xrightarrow{\mu} f_1 \cdot f_2$. 若还有 $f_2^{(n)}(\omega) f_2(\omega) \neq 0$, 对一切 $\omega \in \Omega$, 则

$$\frac{f_1^{(n)}}{f_2^{(n)}} \xrightarrow{\mu} \frac{f_1}{f_2}$$

证明 令 $D := \{(x_1, \cdots, x_m) : Q(x_1, \cdots, x_m) \neq 0\}$, 则 D 为 $\mathbb{R}^n$ 中的开集, 且 $R(x_1, \cdots, x_m)$ 是开集 D 上的连续函数, 因而 (1) 成立. □

下面定理说明 a.e. 收敛与依测度收敛之间的关系.

4. 定理 设 f_n, $n \in \mathbb{N}$, 为可测函数列,

1) 若 $f_n \xrightarrow{\mu} f$, 则存在 $\{f_n\}$ 的子列 $\{f_{n_k}\}$, 使当 $k \to \infty$ 时, $f_{n_k} \xrightarrow{\text{a.e.}} f$.

2) 若 f_n, $n \in \mathbb{N}$, 依测度 μ 相互收敛, 且 (E, d) 完备, 则存在 $\{f_n\}$ 的子列 $\{f_{n_k}\}$ 及一可测函数 f, 使 $f_{n_k} \xrightarrow{\text{a.e.}} f$ $((k \to \infty)$, 且 $f_{n_k} \xrightarrow{\mu} f$.

3) 若 μ 为有限测度, 则 $f_n \xrightarrow{\text{a.e.}} f \Longrightarrow f_n \xrightarrow{\mu} f$.

证明 1) 由于 $f_n \xrightarrow{\mu} f$, 故对一切 $k \in \mathbb{N}$, 存在 $n_k \in \mathbb{N}$ 使当 $n \geq n_k$ 时,

$$\mu\left(\left\{d(f_n, f) \geq \frac{1}{2^k}\right\}\right) < \frac{1}{2^k}$$

而且可取 $n_k\uparrow$, 于是由引理 1.5 知 $f_{n_k}\xrightarrow{\text{a.e.}}f$.

2) 由 $f_n, n\in\mathbb{N}$, 依测度相互收敛立知：对一切 $k\in\mathbb{N}$, 存在 $n_k\in\mathbb{N}$ 使

$$\mu\left(\left\{d(f_{n_k+\nu},f_{n_k})>\frac{1}{2^k}\right\}\right)<\frac{1}{2^k},\qquad \forall\ \nu\geq 1$$

成立，显然不妨设 $n_k\uparrow$. 于是

$$\mu\left(\left\{d(f_{n_{k+1}},f_{n_k})>\frac{1}{2^k}\right\}\right)<\frac{1}{2^k},$$

而 $\sum_{k=1}^{\infty}\frac{1}{2^k}<\infty$, 故由引理 1.6 知 $f_{n_k}, k\in\mathbb{N}$, a.e. 相互收敛，因而由 1.1 的 5) 知存在 f 使 $f_{n_k}\xrightarrow{\text{a.e.}}f$.

其次由引理 1.6 知：对一切 $\varepsilon>0$

$$\mu\left(\bigcup_{\nu=1}^{\infty}\left\{d(f_{n_{k+\nu}},f_{n_k})\geq\varepsilon\right\}\right)\to 0,\quad k\to\infty.$$

令 $A_k:=\bigcup_{\nu=1}^{\infty}\{\omega\in\Omega:\ d(f_{n_{k+\nu}}(\omega),f_{n_k}(\omega))\geq\varepsilon\}$, 则当 $\omega\in A_k^c\cap\{f_{n_k}\to f\}$,

$$d(f(\omega),f_{n_k}(\omega))=\lim_{\nu\to\infty}d(f_{n_{k+\nu}}(\omega),f_{n_k}(\omega))\leq\varepsilon,$$

于是 $\omega\in\{d(f,f_{n_k})>\varepsilon\}$ 时， $\omega\in A_k\cup\{f_{n_k}\not\to f\}$, 因而

$$\mu\left(\left\{d(f,f_{n_k})>\varepsilon\right\}\right)\leq\mu(A_k)\to 0\quad(k\to\infty).$$

即 $f_{n_k}\xrightarrow{\mu}f$.

3) 当测度 μ 有限时，由定理 1.3.(i) 及 $f_n\xrightarrow{\text{a.e.}}f$ 知对任何 $\varepsilon>0$,

$$\mu\left(\left\{d(f_n,f)\geq\varepsilon\right\}\right)\leq\mu\left(\left\{\bigcup_{k=n}^{\infty}\left\{d(f_k,f)\geq\varepsilon\right\}\right)\to 0.$$

故 $f_n\xrightarrow{\mu}f$. □

注 3) 中 μ 为有限测度的假设是不可少的. 例如: 设 μ 是 $\mathbb{R}$ 上的 Lebesgue 测度, $f_n = I_{[-n,n]^c}$, 则 $f_n \to 0$, λ − a.e., 但

$$\mu(|f_n - 0| \geqslant \frac{1}{2}) = \infty \not\to 0.$$

5. 定理 若 (E,d) 完备, f_n, $n \in \mathbb{N}$, 依测度收敛的充要条件是 f_n, $n \in \mathbb{N}$, 依测度相互收敛.

证明 注意对一切 $\varepsilon > 0$, 不等式

$$\mu(d(f,g) \geq \varepsilon) \leq \mu\Big(d(f,h) \geq \frac{\varepsilon}{2}\Big) + \mu\Big(d(h,g) \geq \frac{\varepsilon}{2}\Big) \tag{1}$$

恒成立. 若 $f_n \xrightarrow{\mu} f$, 则对一切 $\varepsilon > 0$, 由 (1)

$$\begin{aligned}\mu\left(\left\{d(f_{n+\nu}, f_n) \geq \varepsilon\right\}\right) &\leq \mu\left(\left\{d(f_{n+\nu}, f) \geq \frac{\varepsilon}{2}\right\}\right) \\ &+ \mu\left(\left\{d(f_n, f) \geq \frac{\varepsilon}{2}\right\}\right) \to 0(\text{对 } \nu \text{ 一致成立!}),\end{aligned}$$

故 f_n, $n \in \mathbb{N}$, 依测度相互收敛.

反之, 若 f_n, $n \in \mathbb{N}$, 依测度相互收敛, 则由定理 4 知存在子序列 $\{f_{n_k}\} \subset \{f_n\}$ 及 f, 使 $f_{n_k} \xrightarrow{\mu} f$ $(k \to \infty)$, 于是对一切 $\varepsilon > 0$, 由 (1),

$$\begin{aligned}\mu(d(f,f_n) \geq \varepsilon) &\leq \mu\Big(d(f,f_{n_k}) \geq \frac{\varepsilon}{2}\Big) + \mu\Big(d(f_n, f_{n_k}) \geq \frac{\varepsilon}{2}\Big) \\ &\to 0(\text{取 } n_k > n, \text{并令} n \to \infty),\end{aligned}$$

即 $f_n \xrightarrow{\mu} f$. □

应用定理 4 可以如下改进控制收敛定理:

6. 推论 设 f_n, $n \in \mathbb{N}$, f 为实 (或复) 值 Borel 可测函数, $f_n \xrightarrow{\mu} f$ $(n \to \infty)$, $|f_n| \leq g$, g 关于 μ 可积, 则 $\int |f_n - f| d\mu \to 0$ $(n \to \infty)$, 因而 $\int f_n d\mu \to \int f\, d\mu$.

证明 由 $\overline{\lim}_{n\to\infty}$ 的定义知存在 $\{f_{n_k}\}\subset\{f_n\}$, 使

$$\lim_{k\to\infty}\int|f_{n_k}-f|\,d\mu=\overline{\lim_{n\to\infty}}\int|f_n-f|\,d\mu.$$

而由 $f_n\xrightarrow{\mu}f$, 知 $|f_{n_k}-f|\xrightarrow{\mu}0$, 于是由定理 4.1) 知存在 $|f_{n_k}-f|$, $k\in\mathbb{N}$, 的子列 $|f_{k'}-f|\xrightarrow{\text{a.e.}}0$. 因而 $f_{k'}\xrightarrow{\text{a.e.}}f$, $|f|\le g$, a.e., 故 $|f_{k'}-f|\le 2g$, a.e., $2g$ 可积, 于是由控制收敛定理知

$$\lim_{k'\to\infty}\int|f_{k'}-f|\,d\mu=0,$$

又 $\int|f_{k'}-f|\,d\mu$ 是 $\int|f_{n_k}-f|\,d\mu$ 的子列, 所以

$$\overline{\lim_{n\to\infty}}\int|f_n-f|\,d\mu=\lim_{k\to\infty}\int|f_{n_k}-f|\,d\mu=\lim_{k'\to\infty}\int|f_{k'}-f|\,d\mu=0.$$

因而

$$\int|f_n-f|\,d\mu\to 0.$$

由积分的线性性质及序性质知

$$\left|\int f_n-\int f\right|=\left|\int(f_n-f)\right|\le\int|f_n-f|\to 0.\quad\square$$

习题

1. 直接证明: 若 $X_n\xrightarrow{\mathbf{P}}X$, $Y_n\xrightarrow{\mathbf{P}}Y$, 则 $X_nY_n\xrightarrow{\mathbf{P}}XY$.

2. 若 $X_n\xrightarrow{\mathbf{P}}X$, f 为有界一致连续函数, 试证 $\mathbf{E}f(X_n)\to\mathbf{E}f(X)$.

3. $X_n\downarrow X$ a.e., 每一 X_n 可积, 且 $\inf\mathbf{E}X_n>-\infty$. 试证 $\mathbf{E}X_n\to\mathbf{E}X$.

§3 L^r 收敛

1. 定义 设 $(\Omega,\mathcal{F},\mu)$ 为测度空间，$r\in(0,\infty)$，用 $L^r(\Omega,\mathcal{F},\mu)$（或简记 $L^r(\mu)$）表示一切使 $\int|f|^r d\mu<\infty$ 的 $\mathcal{F}$ 可测函数类. $\{f_n,\ n\in\mathbb{N},\ f\}\subset L^r(\mu)$，若

$$\int|f_n-f|^r\,d\mu\to 0,$$

则称 $f_n,\ n\in\mathbb{N}$，在 $L^r(\Omega,\mathcal{F},\mu)$ 中收敛于 f. 记作 $f_n\xrightarrow{r}f$（或 $f_n\xrightarrow{L^r(\mu)}f$）.

2. 定理 若 $f_n,n\in\mathbb{N}$，在 $L^r(\mu)$ 中收敛于 f，则 $f_n\xrightarrow{\mu}f$，反之，若存在 $g\in L^r(\mu)$，使 $|f_n|\le g$，则 $f_n\xrightarrow{\mu}f\implies f_n\xrightarrow{L^r(\mu)}f$.

证明 由 Чебышев 不等式：对一切 $\varepsilon>0$，

$$\mu(|f_n-f|\ge\varepsilon)\le\frac{\int|f_n-f|^r}{\varepsilon^r}$$

（因为 $\int|f_n|^r=\int_{|f_n|^r\le\varepsilon}|f_n|^r+\int_{|f_n|\ge\varepsilon}|f_n|^r\ge\varepsilon^r\mu(|f_n|\ge\varepsilon)$）. 故 $f_n\xrightarrow{L^r(\mu)}f\implies f_n\xrightarrow{\mu}f$.

反之，若 $f_n\xrightarrow{\mu}f$、$|f_n|\le g$，则 $|f|\le g$ a.e.（因为存在 $f_{n_k}\xrightarrow{\text{a.e.}}f$），因而 $|f_n-f|^r\xrightarrow{\mu}0$，$|f_n-f|^r\le(2g)^r\in L^r(\mu)$. 故由推论 2.6 知 $\int|f_n-f|^r\to0$. □

3. 推论 若概率空间 $(\Omega,\mathcal{F},\mathbf{P})$ 上的 r.v. 序列 $X_n,n\in\mathbb{N}$，一致有界，则 $X_n\xrightarrow{\mathbf{P}}X\iff X_n\xrightarrow{L^r(\mathbf{P})}X$.

一般有下列

4. 定理 $(\Omega,\mathcal{F},\mathbf{P})$ 上的实（或复）值 r.v. 列 $X_n\xrightarrow{\mathbf{P}}X$ 当且仅当

$$\mathbf{E}\left(\frac{|X_n-X|}{1+|X_n-X|}\right)\to0,\quad(n\to\infty).$$

而且，若将 $(\Omega,\mathcal{F},\mathbf{P})$ 上的 a.e. 相等的 r.v. 看成相同，则

$$\rho(X,Y):=\mathbf{E}\left(\frac{|X-Y|}{1+|X-Y|}\right)$$

是随机变量空间上的距离.

证明 若 $\rho(X,Y)=0$, 则由积分性质及 $|X-Y|/(1+|X-Y|)\geq 0$ 知 $|X-Y|/(1+|X-Y|)=0$ a.e.. 因而 $|X-Y|=0$ a.e. $X=Y$ a.e.. 反之亦然. 由于 $\rho(X,Y)=\rho(X-Y,0)$ 故欲证 ρ 是距离，只需证

$$\mathbf{E}\left(\frac{|X-Y|}{1+|X-Y|}\right)\leq\mathbf{E}\left(\frac{|X|}{1+|X|}\right)+\mathbf{E}\left(\frac{|Y|}{1+|Y|}\right).$$

实际上，只需证

$$\frac{|x-y|}{1+|x-y|}\leq\frac{|x|}{1+|x|}+\frac{|y|}{1+|y|},$$

由 $x/(1+x)\uparrow$ $(x\geq 0)$ (因为 $(1+x)/x=1+1/x\downarrow$), 知

$$\begin{aligned}\frac{|x|}{1+|x|}+\frac{|y|}{1+|y|}&\geq\frac{|x|+|y|+|xy|}{1+|x|+|y|+|xy|}\\&\geq\frac{|x|+|y|}{1+|x|+|y|}\geq\frac{|x-y|}{1+|x-y|}.\end{aligned}$$

故 ρ 是距离.

今往证定理的第一部分.

$$\begin{aligned}&X_n\xrightarrow{\mathbf{P}}X\\ \Longleftrightarrow&\ \forall\,\varepsilon>0,\quad\mathbf{P}\left(\frac{|X_n-X|}{1+|X_n-X|}\geq\frac{\varepsilon}{1+\varepsilon}\right)\\&=\mathbf{P}(|X_n-X|\geq\varepsilon)\to 0\\ \Longleftrightarrow&\ \frac{|X_n-X|}{1+|X_n-X|}\xrightarrow{\mathbf{P}}0\\ \Longleftrightarrow&\ \mathbf{E}\left(\frac{|X_n-X|}{1+|X_n-X|}\right)\to 0.\quad\square\end{aligned}$$

和距离空间中的证明一样 (见 2.1.3 例 2), 由初等不等式 (见习题 2.1.1)

$$|ab| \leq \frac{|a|^p}{p} + \frac{|b|^q}{q}, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad p > 1,$$

可得

5. Hölder不等式: 设 $X \in L^p(\mathbf{P})$, $Y \in L^q(\mathbf{P})$, $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, 则

$$\mathbf{E}|XY| \leq [\mathbf{E}|X|^p]^{1/p}[\mathbf{E}|Y|^q]^{1/q}$$

(当 $p = q = 2$ 时, 即为 Schwartz 不等式).

由此易得 (方法与例 2.1.3 例 2 的证明一样!)

6. Minkowski不等式: 若 $r \geq 1$, $X, Y \in L^r(\mathbf{P})$, 则

$$[\mathbf{E}|X+Y|^r]^{1/r} \leq \mathbf{E}^{1/r}|X|^r + \mathbf{E}^{1/r}|Y|^r.$$

于是令

$$d(X,Y) = \mathbf{E}^{1/r}|X-Y|^r, \qquad 当 X, Y \in L^r(\mathbf{P}),$$

且认为 $\mathbf{P}(X=Y)=1$ 时, X, Y 相同, 则 $d(X,Y)$, $X, Y \in L^r(\mathbf{P})$ 是 $L^r(\mathbf{P})$ 中的距离, 因而有

7. 引理　设 $r \geq 1$, $X_n \xrightarrow{r} X \implies \mathbf{E}|X_n|^r \to \mathbf{E}|X|^r$.

证明

$$\begin{aligned} \left|\mathbf{E}^{1/r}|X_n|^r - \mathbf{E}^{1/r}|X|^r\right| &= |d(X_n,0) - d(X,0)| \\ &\leqslant d(X_n, X) = \mathbf{E}^{1/r}|X_n - X|^r \to 0. \quad \square \end{aligned}$$

8. 引理　$\mathbf{E}^{1/r}|X|^r$, $r > 0$, 对 r 不降. 且 $\lim_{r\uparrow p}\mathbf{E}^{1/r}|X|^r = \mathbf{E}^{1/p}|X|^p$.

证明 设 $0 < r < p$, 分别用 $|X|^r$, 1^p, p/r, $p/(p-r)$ 代替 Hölder 不等式中的 X, Y, p, q, 则

$$\mathbf{E}|X|^r \leq \mathbf{E}^{r/p}(|X|^r)^{p/r} = \mathbf{E}^{r/p}|X|^p,$$

于是

$$\mathbf{E}^{1/r}|X|^r \leq \mathbf{E}^{1/p}|X|^p.$$

对 $\mathbf{E}[|X|^r I_{\{|X|\leq 1\}}]$ 和 $\mathbf{E}\left[|X|^r I_{\{|X|>1\}}\right]$ 分别应用控制收敛定理及单调收敛定理即得后一结论. □

为讨论数学期望的重要性质，我们给出凸函数的定义并讨论其性质.

9. 定义 若 φ 为区间 (a,b) 上的实函数，如果对一切 $a < x < y < b$ 及一切 $\lambda \in (0,1)$, 有

$$\varphi(\lambda x + (1-\lambda)y) \leqslant \lambda\varphi(x) + (1-\lambda)\varphi(y). \tag{1}$$

则称 φ 为区间 (a,b) 上的凸函数. $\mathbb{R}$ 中的凸函数简称为凸函数.

10. 例 若 $\varphi(r) = \log \mathbf{E}|X|^r$ 在 (a,b) 上有定义，则 φ 为 (a,b) 上的凸函数.

证明 令 $\lambda \in (0,1)$, 则由 Hölder 不等式将其中的 X, Y, $1/p$, $1/q$ 分别代以 $|X|^{\lambda r}$, $|Y|^{(1-\lambda)s}$, λ, $1-\lambda$,

$$\begin{aligned}\varphi(\lambda r + (1-\lambda)s) &= \log \mathbf{E}|X|^{\lambda r} \cdot |X|^{(1-\lambda)s} \\ &\leq \log \mathbf{E}^{\lambda}|X|^r \cdot \mathbf{E}^{(1-\lambda)}|X|^s \\ &= \lambda \log \mathbf{E}|X|^r + (1-\lambda)\log \mathbf{E}|X|^s \\ &= \lambda\varphi(r) + (1-\lambda)\varphi(s). \quad \square\end{aligned}$$

11. 引理 若 φ 为区间 (a,b) 上的凸函数，则对一切 $x \in (a,b)$, 存在有限的左右导数 $\varphi'_\ell(x), \varphi'_r(x)$, 且对一切 $a < x < y < b$,

$$\varphi'_\ell(x) \leqslant \varphi'_r(x) \leqslant \varphi'_\ell(y) \leqslant \varphi'_r(y). \tag{1}$$

证明 几何地说，由凸函数的定义知，对一切 $a \leqslant x \leqslant y \leqslant b$，$\varphi$ 在 $[x,y]$ 上的图象都在联结 $(x,\varphi(x))$，$(y,\varphi(y))$ 的直线之下. 于是，对一切 $a < x_1 < x < x_2 < b$，有

$$\frac{\varphi(x)-\varphi(x_1)}{x-x_1} \le \frac{\varphi(x_2)-\varphi(x_1)}{x_2-x_1} \le \frac{\varphi(x_2)-\varphi(x)}{x_2-x}, \tag{2}$$

事实上，此时存在 $\lambda \in (0,1)$ 使 $x = \lambda x_1 + (1-\lambda)x_2$，由定义 9 知

$$\begin{aligned}\frac{\varphi(x)-\varphi(x_1)}{x-x_1} &\le \frac{\lambda\varphi(x_1)+(1-\lambda)\varphi(x_2)-\varphi(x_1)}{\lambda x_1+(1-\lambda)x_2-x_1}\\ &= \frac{\varphi(x_2)-\varphi(x_1)}{x_2-x_1},\end{aligned}$$

同法可证另一不等式.

其次由 (2) 知对一切 $x_1' \in (x_1,x)$，$x_2' \in (x,x_2)$，有

$$\begin{aligned}\frac{\varphi(x_1)-\varphi(x)}{x_1-x} &\le \frac{\varphi(x_1')-\varphi(x)}{x_1'-x}\\ \le \frac{\varphi(x_2')-\varphi(x)}{x_2'-x} &\leqslant \frac{\varphi(x_2)-\varphi(x)}{x_2-x},\end{aligned} \tag{3}$$

在 (3) 中令 $x_1' \uparrow x$，即知 $\varphi_\ell'(x)$ 存在且

$$\varphi_\ell'(x) \le \frac{\varphi(x_2')-\varphi(x)}{x_2'-x}, \ \forall\ x_2' > x.$$

再令 $x_2' \downarrow x$ 即知 $\varphi_r'(x)$ 存在且

$$\varphi_r'(x) \ge \frac{\varphi(x_1')-\varphi(x)}{x_1'-x}, \ \forall\, x_1' < x$$

成立，于是 $\varphi_\ell'(x) \leqslant \varphi_r'(x)$. 至于 $\varphi_r'(x) \leqslant \varphi_\ell'(y)$，$x < y$ 的正确性可由 (2) 及左，右导数的存在性得到. 故得 (1) 式. □

12. Jessen 不等式. 若 φ 为 $\mathbb{R}$ 上的凸函数，X，$\varphi(X)$ 都是可积 r.v.，则

$$\varphi(\mathbf{E}(X)) \le \mathbf{E}\varphi(X).$$

证明 由引理 11 知

$$\text{(1)}\qquad \begin{aligned}&\frac{\varphi(y)-\varphi(x)}{y-x}\downarrow\varphi_r'(x)\quad(y\downarrow x)\\&\geq\varphi_l'(x)\uparrow\frac{\varphi(y)-\varphi(x)}{y-x}\quad(y\uparrow x)\end{aligned}$$

于是得知 φ 为 $\mathbb{R}$ 上的连续函数，因而 $\varphi(X)$ 为 r.v.，且

$$\varphi(y)-\varphi(x)\geq\varphi_r'(x)(y-x),\qquad\forall y\in\mathbb{R},$$

令 $y=X$, $x=\mathbf{E}X$, 则得

$$\varphi(X)-\varphi(\mathbf{E}X)\geq\varphi'(\mathbf{E}X)(X-\mathbf{E}X),$$

两边取积分，由积分单调性即得 Jessen 不等式. □

13. 引理 (基本不等式) 设 φ 是 $\mathbb{R}$ 上取正值的偶函数，且在 $(0,\infty)$ 上不降， X 是一 r.v. 且 $\mathbf{E}\varphi(X)<\infty$, 则对一切 $u>0$

$$\text{(1)}\qquad \frac{\mathbf{E}\varphi(X)-\varphi(u)}{\text{a.s.}\sup\varphi(X)}\leq\mathbf{P}(|X|\geq u)\leq\frac{\mathbf{E}\varphi(X)}{\varphi(u)}$$

若 φ 是 $\mathbb{R}$ 上严格正值增函数，则对一切 $u\in\mathbb{R}$

$$\text{(2)}\qquad \frac{\mathbf{E}\varphi(X)-\varphi(u)}{\text{a.s.}\sup\varphi(X)}\leq\mathbf{P}(X\geq u)\leq\frac{\mathbf{E}\varphi(X)}{\varphi(u)}$$

其中 $\text{a.s.}\sup\varphi(X)=\inf\{c:\ \mathbf{P}(\varphi(X)\geqslant c)=0\}$.

证明 显然 φ 为实可测函数，所以 $\varphi(X)$ 是 r.v.，而

$$\mathbf{E}\varphi(X)=\int_{\{|X|\geq u\}}\varphi(X)\,d\mathbf{P}+\int_{\{|X|<u\}}\varphi(X)\,d\mathbf{P};$$

$$\begin{aligned}\varphi(u)\mathbf{P}(|X|\geq u)&\leq\int_{\{|X|\geq u\}}\varphi(X)\,d\mathbf{P}\\&\leq\text{a.s.}\sup\varphi(X)\cdot\mathbf{P}(|X|\geq u),\end{aligned}$$

故

$$\mathbf{P}(|X| \geq u) \leq \frac{\mathbf{E}\varphi(X)}{\varphi(u)},$$

$$\begin{aligned}\mathbf{P}(|X| \geq u) &\geq \frac{\int_{\{|X| \geq u\}} \varphi(x)\, d\mathbf{P}}{\text{a.s.}\sup \varphi(X)} \\ &= \frac{\mathbf{E}\varphi(X) - \int_{\{|X| < u\}} \varphi(X)\, d\mathbf{P}}{\text{a.s.}\sup \varphi(X)} \geqslant \frac{\mathbf{E}\varphi(X) - \varphi(u)}{\text{a.s.}\sup \varphi(X)}.\end{aligned}$$

同理可证 (2). □

下面讨论 $L^r(\mathbf{P})$ 收敛及矩收敛的有关条件.

14. 引理 若 $X_n \xrightarrow{\text{a.e.}} X$, 则对一切 $r > 0$,

$$\mathbf{E}|X|^r \leq \varliminf_{n\to\infty} \mathbf{E}|X_n|^r.$$

证明 由 Fatou 引理知

$$\mathbf{E}|X|^r = \mathbf{E}\left(\lim_{n\to\infty} |X_n|^r\right) \leq \varliminf_{n\to\infty} \mathbf{E}|X_n|^r. \quad \square$$

现在我们引进重要的一致可积的概念.

15. 定义 设 $\{X_t:\ t \in T\}$ 为任一族 r.v. , 如果

$$\lim_{C\to\infty} \int_{\{|X_t| \geq C\}} |X_t|\, d\mathbf{P} = 0$$

对 $\{X_t:\ t \in T\}$ 一致成立 (即 $\lim_{C\to\infty} \sup_{t\in T} \int_{\{|X_t|\geq C\}} |X_t|\, d\mathbf{P} = 0$), 则称 $\{X_t:\ t \in T\}$ **一致可积**.

16. 引理 $\{X_t:\ t \in T\}$ 一致可积的充要条件是下列两条件满足:

1) $\sup_{t\in T} \mathbf{E}|X_t| < \infty$;

2) 对一切 $\varepsilon > 0$, 存在 $\delta(\varepsilon) > 0$, 使对一切 $A \in \mathfrak{F}$,

$$\mathbf{P}(A) < \delta(\varepsilon) \implies \int_A |X_t|\, d\mathbf{P} < \varepsilon, \qquad \forall\, t \in T.$$

此条件称为 $\{X_t:\ t\in T\}$**积分一致绝对连续**，也可表成

$$\lim_{\mathbf{P}(A)\to 0}\sup_{t\in T}\int_A |X_t|\,d\mathbf{P}=0.$$

证明 若 $\{X_t:\ t\in T\}$ 一致可积，则

$$\begin{aligned}\mathbf{E}|X_t| &= \int_{\{|X_t|\ge C\}}|X_t|\,d\mathbf{P}+\int_{\{|X_t|<C\}}|X_t|\,d\mathbf{P}\\ &\le C+\int_{\{|X_t|\ge C\}}|X_t|\,d\mathbf{P},\end{aligned}$$

因而 1) 成立，类似方法有

$$\begin{aligned}\int_A |X_t|\,d\mathbf{P} &= \int_{A\cap\{|X_t|\ge C\}}|X_t|\,d\mathbf{P}+\int_{A\cap\{|X_t|<C\}}|X_t|\,d\mathbf{P}\\ &\le \int_{\{|X_t|\ge C\}}|X_t|\,d\mathbf{P}+C\mathbf{P}(A).\end{aligned}$$

于是 2) 成立.

反之，设 1), 2) 成立．由基本不等式知

$$\mathbf{P}(|X_t|\ge C)\le\frac{\mathbf{E}|X_t|}{C}\le\frac{\sup_{t\in T}\mathbf{E}|X_t|}{C}.$$

对一切 $\varepsilon>0$, 由 2) 知存在 $\delta=\delta(\varepsilon)>0$ 使

$$\mathbf{P}(|X_t|\ge C)<\delta\implies\int_{\{|X_t|\ge C\}}|X_t|\,d\mathbf{P}<\varepsilon.$$

而由 1) 知存在 $C>(1/\delta)\sup_{t\in T}\mathbf{E}|X_t|$, 此时

$$\sup_{t\in T}\int_{\{|X_t|\ge C\}}|X_t|\,d\mathbf{P}<\varepsilon,$$

所以 $\{X_t:\ t\in T\}$ 一致可积. □

17. 引理 (Fatou 引理及控制收敛定理的推广) (I) 设 U_n, V_n 可积，且 $U_n\xrightarrow{\text{a.e.}}U$, $V_n\xrightarrow{\text{a.e.}}V$, $\int U_n\to\int U$ 有限， $\int V_n\to\int V$ 有限.

1) 若 $U_n \le X_n$, 对一切 $n \in \mathbb{N}$, 则

$$\int \varliminf_{n\to\infty} X_n \le \varliminf_{n\to\infty} \int X_n$$

2) 若 $X_n \le V_n$, 对一切 $n \in \mathbb{N}$, 则

$$\varlimsup_{n\to\infty} \int X_n \le \int \varlimsup_{n\to\infty} X_n$$

(II) 设 $|X_n| \le U_n$ 而 U_n 可积 $U_n \xrightarrow{\mu} U$ 且 $\int U_n \to \int U$ 有限, 则当 $X_n \xrightarrow{\mu} X$ (或 $X_n \xrightarrow{\text{a.e.}} X$) 时, 有

$$\int |X_n - X| \to 0, \qquad \text{因而} \qquad \int X_n \to \int X.$$

(I) 即习题 5.4.1, (II) 将习题 5.4.2 中的 a.e. 收敛换作依测度收敛. 证明留作习题.

18. 定理 设 $r \in (0,\infty)$, $X_n \in L^r(\mathbf{P})$, $n \in \mathbb{N}$, 且 $X_n \xrightarrow{\mathbf{P}} X$, 则下述三命题等价:

1) $\{|X_n|^r: n \in \mathbb{N}\}$ 一致可积;

2) $X_n \xrightarrow{L^r(\mathbf{P})} X$;

3) $\mathbf{E}|X_n|^r \to \mathbf{E}|X|^r < \infty$.

证明 1) $\Longrightarrow$ 2). 存在 $\{X_{n_k}: k \in \mathbb{N}\}$ 使 $X_{n_k} \xrightarrow{\text{a.s.}} X$, 于是由 1) 及引理 14 知

$$\mathbf{E}|X|^r \le \varlimsup_{n\to\infty} \mathbf{E}|X_n|^r < \infty,$$

即 $X \in L^r(\mathbf{P})$, 再由 $\{|X_n|^r: n \in \mathbb{N}\}$ 一致可积, 引理 16.2) 及

$$|X_n - X|^r \le 2^r(|X_n|^r + |X|^r),$$

(因为 $|X_n - X|^r \le \left[2(|X_n| \vee |X|)\right]^r = 2^r(|X_n|^r \vee |X|^r) \le 2^r(|X_n|^r + |X|^r)$) 知 $|X_n - X|^r n \in \mathbb{N}$ 积分一致绝对连续, 于是对一切 $\varepsilon > 0$,

$$\mathbf{E}|X_n - X|^r \le \int_{\{|X_n - X| \ge \varepsilon\}} |X_n - X|^r \, d\mathbf{P} + \varepsilon^r,$$

而由假设知 $\mathbf{P}(|X_n - X| \geq \varepsilon) \to 0$ $(n \to \infty)$ 及 $|X_n - X|^r$ 一致绝对连续知

$$\overline{\lim_{n\to\infty}} \mathbf{E}|X_n - X|^r \leq 2\varepsilon^r.$$

令 $\varepsilon \to 0$ 即得 2).

2) $\Longrightarrow$ 3). 对于 $r \geq 1$ 的情形，即引理 7. 对于 $r \in (0,1)$, 有不等式

$$|x+y|^r \leq |x|^r + |y|^r, \qquad x, y \in \mathbb{R}$$

由此即得

$$\mathbf{E}|X_n|^r - \mathbf{E}|X_n - X|^r \leq \mathbf{E}|X|^r \leq \mathbf{E}|X_n|^r + \mathbf{E}|X - X_n|^r,$$

因而由 2)

$$|\mathbf{E}|X_n|^r - \mathbf{E}|X|^r| \leq \mathbf{E}|X_n - X|^r \to 0.$$

3) $\Longrightarrow$ 2). 为此令 $U_n := 2^r(|X_n|^r + |X|^r)$, 则

$$|X_n - X|^r \leq U_n, \quad 且 \quad U_n \xrightarrow{\mathbf{P}} 2^{r+1}|X|^r,$$

后一结论用到定理 2.2.2) (也可应用有关不等式直接证明). 由 3)

$$\int U_n \, d\mathbf{P} = 2^r \int |X_n| \, d\mathbf{P} + 2^r \mathbf{E}|X|^r \to 2^{r+1}\mathbf{E}|X|^r,$$

再由引理 17.(II) 即得 2) (即 $\int |X_n - X|^r \, d\mathbf{P} \to 0$). 由 2) 易得 1) (分 $r \geq 1$, $r \in (0,1)$ 两种情形讨论).

19. 推论 1) 若 $X_n \xrightarrow{r} X$, 则 $X_n \xrightarrow{r'} X$, 对一切 $r' \in (0, r)$;

2) 若 $\sup \mathbf{E}|X_n|^r = c < \infty$, 则

$$X_n \xrightarrow{\mathbf{P}} X \Longrightarrow X_n \xrightarrow{r'} X, \quad r \in (0, r);$$

3) 若 $|X_n| \leq Y \in L^r(\mathbf{P})$, 对一切 n 成立，则

$$X_n \xrightarrow{\mathbf{P}} X \Longrightarrow X_n \xrightarrow{r} X \in L^r(\mathbf{P}).$$

证明 1) 由引理 8 知

$$\mathbf{E}^{1/r'}|X_n - X|^{r'} \le \mathbf{E}^{1/r}|X_n - X|^r, \quad \forall r' < r,$$

即得 1).

2) 由引理 8 知对一切 $r' < r$,

$$\sup_n \mathbf{E}|X_n|^{r'} \le \sup_n (\mathbf{E}|X_n|^r)^{r'/r} = c^{r'/r} < \infty.$$

其次

$$\int_A |X_n|^{r'}\, d\mathbf{P} = \int_{A\cap\{|X_n|\ge a\}} |X_n|^{r'}\, d\mathbf{P} + \int_{A\cap\{|X_n|<a\}} |X_n|^{r'}\, d\mathbf{P}$$
$$\le \int_{A\cap\{|X_n|\ge a\}} \frac{|X_n|^r}{a^{r-r'}}\, d\mathbf{P} + a^{r'}\mathbf{P}(A) \le ca^{r'-r} + a^{r'}\mathbf{P}(A).$$

对一切 $\varepsilon > 0$, 先取定 a 充分大使 $ca^{r'-r} < \varepsilon/2$, 于是当 $\mathbf{P}(A)$ 充分小使 $a^{r'}\mathbf{P}(A) < \varepsilon/2$, 因而当 $\mathbf{P}(A)$ 充分小时, 对一切 $n \in \mathbb{N}$, $\int_A |X_n|^{r'}\, d\mathbf{P} < \varepsilon$. 由引理 18 知 $|X_n|^{r'}, n \in \mathbb{N}$ 对 $\mathbf{P}$ 一致可积. 故由定理 1 知 $X_n \xrightarrow{r'} X$. 2) 获证.

3) 由于 $|X_n - X| \le 2Y$ 对一切 n 成立, 且 $|X_n - X|^r \xrightarrow{\mathbf{P}} 0$, 所以由控制收敛定理的改进 (推论 2.6) 知 $\mathbf{E}|X_n - X|^r \to 0$. □

习题

1. 证明 Hölder 不等式及 Minkowski 不等式.

2. 试完成定理 18 中的 2) $\Longrightarrow$ 1).

3. 设 $r \in (0, \infty)$, 则 $X_n \xrightarrow{r} X$ 当且仅当 $X_n \xrightarrow{\mathbf{P}} X$ 及下列两条件之一成立:

1) $\mathbf{E}|X_n|^r \to \mathbf{E}|X|^r < \infty$;

2) $|X_n|^r$ 一致可积.

4. 若 $\sup_n |X_n| \in L^r(\mathbf{P})$, 且 $X_n \xrightarrow{\text{a.e.}} X$, 则 $X \in L^r(\mathbf{P})$, 且 $X_n \xrightarrow{r} X$.

5. 证明引理 17 的 (II).

§4 条件期望的进一步性质

下面应用与上节的定理类似的方法及条件期望的平滑性质证明关于条件期望的一些进一步的性质:

1. 定理 条件期望有下列进一步的期望性质: 设下列各结论中随机变量的期望都有限, 则

(I)(Hölder 不等式). 若 $r>1$, $s>1$, $\frac{1}{r}+\frac{1}{s}=1$, 则

$$\mathbf{E}[|XY||\mathcal{C}]\le[\mathbf{E}[|X|^r|\mathcal{C}]]^{\frac{1}{r}}[\mathbf{E}[|Y|^s|\mathcal{C}]]^{\frac{1}{s}},\ \mathbf{P}_{\mathcal{C}}-\text{a.e.}.$$

(II)(Minkowski 不等式). 若 $r\ge 1$, 则

$$[\mathbf{E}[|X+Y|^r|\mathcal{C}]]^{\frac{1}{r}}\le[\mathbf{E}[|X|^r|\mathcal{C}]]^{\frac{1}{r}}+[\mathbf{E}[|Y|^r|\mathcal{C}]]^{\frac{1}{r}},\ \mathbf{P}_{\mathcal{C}}-\text{a.e.};$$

若 $r<1$, 则

$$\mathbf{E}[|X+Y|^r|\mathcal{C}]\le\mathbf{E}[|X|^r|\mathcal{C}]+\mathbf{E}[|Y|^r|\mathcal{C}],\ \mathbf{P}_{\mathcal{C}}-\text{a.e.}.$$

(III)(Jessen 不等式). 若 φ 是 $\mathbb{R}$ 上的凸函数, 且 X 和 $\varphi(X)$ 都是可积随机变量, 则

$$\varphi(\mathbf{E}[X|\mathcal{C}])\le\mathbf{E}[\varphi(X)|\mathcal{C}],\ \mathbf{P}_{\mathcal{C}}-\text{a.e.}.$$

(IV) $[\mathbf{E}[|X|^r|\mathcal{C}]]^{\frac{1}{r}}$, $\mathbf{P}_{\mathcal{C}}$–a.e. 对 $r>0$ 不降.

证明 (I) 的证明: 与积分的 Hölder 不等式的证明原则上相同. 由假设 $X\in L^r(\Omega,\mathcal{F},\mathbf{P})$, $Y\in L^s(\Omega,\mathcal{F},\mathbf{P})$, 因而由 Hölder 不等式知 XY 可积. 令 $B_1:=\{\omega:\ \mathbf{E}[|X|^r|\mathcal{C}]=0\}$, $B_2:=\{\omega:\mathbf{E}[|Y|^s|\mathcal{C}]=0\}$, 于是 B_1, $B_2\in\mathcal{C}$. 且

$$\int_{B_1}|X|^r\,d\mathbf{P}=\int_{B_1}\mathbf{E}[|X|^r|\mathcal{C}]d\mathbf{P}=0,$$

因此 $I_{B_1}|X|^r = 0$, $\mathbf{P}-$a.e.. 同理可证 $I_{B_2}|Y|^s =,0$ $\mathbf{P}-$a.e., 因而 $I_{B_1\cup B_2}|XY| = 0$, $\mathbf{P}-$a.e.. 于是由定理 7.4.3 及 $B_1 \cup B_2 \in \mathcal{C}$ 知

$$I_{B_1\cup B_2}\mathbf{E}[|XY||\mathcal{C}] = \mathbf{E}[I_{B_1\cup B_2}|XY||\mathcal{C}] = 0, \ \mathbf{P}_{\mathcal{C}} - \text{a.e.},$$

故 Hölder 不等式在 $B_1 \cup B_2$ 上成立.

在 $(B_1 \cup B_2)^c = B_1^c \cap B_2^c$ 上 $\mathbf{E}[|X|^r|\mathcal{C}] \neq 0$, $\mathbf{E}[|Y|^s|\mathcal{C}] \neq 0$, 于是将

$$a := \frac{|X|^r}{\mathbf{E}[|X|^r|\mathcal{C}]}, \quad b := \frac{|Y|^s}{\mathbf{E}[|Y|^s|\mathcal{C}]},$$

代入初等不等式

$$a^{\frac{1}{r}}b^{\frac{1}{s}} \leq \frac{a}{r} + \frac{b}{s}, \quad a \geq 0, b \geq 0,$$

即得

$$\frac{|XY|}{[\mathbf{E}[|X|^r|\mathcal{C}]]^{\frac{1}{r}}[\mathbf{E}[|Y|^s|\mathcal{C}]]^{\frac{1}{s}}} \leq \frac{1}{r}\frac{|X|^r}{\mathbf{E}[|X|^r|\mathcal{C}]} + \frac{1}{s}\frac{|Y|^s}{\mathbf{E}[|Y|^s|\mathcal{C}]}.$$

两边乘以 $I_{B_1^c\cap B_2^c}$, 并取在 $\mathcal{C}$ 之下的条件期望, 注意到 $[\mathbf{E}[|X|^r|\mathcal{C}]]^{\frac{1}{r}}$, $[\mathbf{E}[|Y|^s|\mathcal{C}]]^{\frac{1}{s}}$ 及 $I_{B_1^c\cap B_2^c}$ 的 $\mathcal{C}-$ 可测性, 并应用定理 7.4.3 及 7.4.1 的 (III) 即得

$$I_{B_1^c\cap B_2^c}\frac{\mathbf{E}[|XY||\mathcal{C}]}{[\mathbf{E}[|X|^r|\mathcal{C}]]^{\frac{1}{r}}[\mathbf{E}[|Y|^s|\mathcal{C}]]^{\frac{1}{s}}} \leq I_{B_1^c\cap B_2^c}, \ \mathbf{P}_{\mathcal{C}} - \text{a.e.}.$$

即在 $B_1^c \cap B_2^c$ 上 Hölder 不等式仍然成立. 故 (I) 获证.

(II) 的证明：由假设 $\mathbf{E}|X|^r$, $\mathbf{E}|Y|^r$ 都有限. 当 $r > 1$ 时，其证明和积分的 Minkowski 不等式的证明一样. 当 $r < 1$ 时，有

$$|X+Y|^r \leq |X|^r + |Y|^r$$

两边取 $\mathcal{C}$ 之下的条件期望，即得 (II) 的第二个结论.

(III) 的证明：在 Jessen 不等式的证明中，我们实际已经证明：对任何 $x, y \in \mathbb{R}$, $y < x$, 有

$$\varphi_r'(x) \geq \varphi_\ell'(x) \geq \frac{\varphi(x) - \varphi(y)}{x - y} \geq \varphi_r'(y).$$

于是 φ'_r 不降，且

$$\text{(1)}\qquad \varphi(X)-\varphi(\mathbf{E}[X|\mathcal{C}])\geq\varphi'_r(\mathbf{E}[X|\mathcal{C}])(X-\mathbf{E}[X|\mathcal{C}]).$$

若 (1) 式右边可积，则对两边取在给定 $\sigma-$ 代数 $\mathcal{C}$ 下的条件期望即得 (III) (为什么？). 若不可积，则对任何 $a>0$, 令 $A:=\{\omega:|\mathbf{E}[X|\mathcal{C}]|\leqslant a\}$. 由 φ'_r 不降及 $|\mathbf{E}[XI_A|\mathcal{C}]|\leqslant a$ 知

$$\varphi'(-a)\leqslant\varphi'_r(\mathbf{E}[XI_A|\mathcal{C}]\leqslant\varphi'_r(a).$$

于是将 (1) 中的 X 换以 XI_A 后，右边就可积了. 此时再对两边取在给定 $\sigma-$ 代数 $\mathcal{C}$ 下的条件期望, 由平滑性质即得 (注意 $\varphi(XI_A)=\varphi(X)I_A+\varphi(0)I_{A^c}$)

$$\begin{aligned}0&\leq\mathbf{E}[\varphi(XI_A)|\mathcal{C}]-\varphi(\mathbf{E}[XI_A|\mathcal{C}])\\&=\mathbf{E}[\varphi(X)|\mathcal{C}]I_A+\varphi(0)I_{A^c}-\varphi(\mathbf{E}[X|\mathcal{C}]I_A)\\&=\mathbf{E}[\varphi(X)|\mathcal{C}]I_A-\varphi(\mathbf{E}[X|\mathcal{C}])I_A\\&\longrightarrow\mathbf{E}[\varphi(X)|\mathcal{C}]-\varphi(\mathbf{E}[X|\mathcal{C}]),\ (a\to\infty),\end{aligned}$$

即 (III) 获证.

(IV) 的证明与无条件的期望的相应结论的证明一样. □

2. 定理 $X_n\xrightarrow{r}X$, $r\geq1$,

$$\mathbf{E}[X_n|\mathcal{C}]\xrightarrow{r}\mathbf{E}[X|\mathcal{C}].$$

证明 由假设知 $\mathbf{E}|X_n|$, $\mathbf{E}|X|$, $\mathbf{E}|X_n-X|^r$ 有限，因而由定理 7.4.5 及上述定理的 (IV) 知

$$\begin{aligned}|\mathbf{E}[X_n|\mathcal{C}]-\mathbf{E}[X|\mathcal{C}]|&\leq\mathbf{E}[|X_n-X||\mathcal{C}]\\&\leq[\mathbf{E}[|X_n-X|^r|\mathcal{C}]]^{\frac{1}{r}},\ \mathbf{P}_{\mathcal{C}}-\text{a.e.}.\end{aligned}$$

由此及假设知

$$\begin{aligned}\mathbf{E}|\mathbf{E}[X_n|\mathcal{C}]-\mathbf{E}[X|\mathcal{C}]|^r&\leq\mathbf{E}[\mathbf{E}[|X_n-X|^r|\mathcal{C}]]\\&=\mathbf{E}|X_n-X|^r\to0.\end{aligned}$$

故定理获证. □

最后，我们证明条件期望在非线性最佳均方逼近问题上的一个应用.

3. 定理 若 $X \in L^2(\Omega, \mathcal{F}, \mathbf{P})$, $\mathcal{C}$ 是 $\mathcal{F}$ 的子 $\sigma-$ 代数，则 $\mathbf{E}[X|\mathcal{C}] \in L^2(\Omega, \mathcal{C}, \mathbf{P}_\mathcal{C})$, 且 $\mathbf{E}[X|\mathcal{C}]$ 是 X 在 $L^2(\Omega, \mathcal{C}, \mathbf{P}_\mathcal{C})$ 中的 **最佳均方逼近**, 即对任何 $f \in L^2(\Omega, \mathcal{C}, \mathbf{P}_\mathcal{C})$ 来说,

$$\mathbf{E}|X - \mathbf{E}[X|\mathcal{C}]|^2 \leq \mathbf{E}|X - f|^2, \tag{1}$$

事实上，下列更强的不等式成立：

$$\mathbf{E}[|X - \mathbf{E}[X|\mathcal{C}]|^2|\mathcal{C}] \leq \mathbf{E}[|X - f|^2|\mathcal{C}], \ \mathbf{P}_\mathcal{C} - \text{a.e.}. \tag{2}$$

证明 只需证后一结论，因为对它取数学期望即得前一结论. 我们还注意在定理 2 的证明中已证 $\mathbf{E}[X|\mathcal{C}] \in L^2(\Omega, \mathcal{C}, \mathbf{P}_\mathcal{C})$. 现在进行定理的证明.

因为 $|a+b|^2 = |a|^2+|b|^2+a\bar{b}+\bar{a}b$, 所以对任何 $f \in L^2(\Omega, \mathcal{C}, \mathbf{P}_\mathcal{C})$,

$$\begin{aligned}
\mathbf{E}[|X - f|^2|\mathcal{C}] &= \mathbf{E}[|(X - \mathbf{E}[X|\mathcal{C}]) + (\mathbf{E}[X|\mathcal{C}] - f)|^2|\mathcal{C}] \\
&= \mathbf{E}[|X - \mathbf{E}[X|\mathcal{C}]|^2|\mathcal{C}] + \mathbf{E}[|\mathbf{E}[X|\mathcal{C}] - f|^2|\mathcal{C}] \\
&\quad + \mathbf{E}[(X - \mathbf{E}[X|\mathcal{C}])\overline{(\mathbf{E}[X|\mathcal{C}] - f)}|\mathcal{C}] \\
&\quad + \mathbf{E}[\overline{(X - \mathbf{E}[X|\mathcal{C}])}(\mathbf{E}[X|\mathcal{C}] - f)|\mathcal{C}].
\end{aligned} \tag{3}$$

由条件期望的平滑性质及 $\mathbf{E}[X|\mathcal{C}] - f$ 的 $\mathcal{C}-$ 可测性知 (3) 式右端的第三项

$$\begin{aligned}
&\mathbf{E}[(X - \mathbf{E}[X|\mathcal{C}])\overline{(\mathbf{E}[X|\mathcal{C}] - f)}|\mathcal{C}] \\
&= \overline{(\mathbf{E}[X|\mathcal{C}] - f)}\mathbf{E}[X - \mathbf{E}[X|\mathcal{C}]|\mathcal{C}] = 0.
\end{aligned}$$

而 (3) 式右边的第四项为第三项的共轭数，所以也等于零. 还有 (3) 式右边的第二项非负. 于是由 (3) 即得 (2). □

4. 注 事实上，常常遇到这样的问题：设 $X, X_1, \cdots, X_n$ 是二阶矩存在的随机变量. 希望用 $X_1, \cdots, X_n$ 的一个可测函数

$f(X_1,\cdots,X_n)$ 去最佳逼近 X(通常称为回归问题). 若"最佳逼近"的意义理解为"最佳均方逼近", 则由定理 3 及定理 7.2.5 知

$$f(X_1,\cdots,X_n)=\mathbf{E}[X|\sigma(X_1,\cdots,X_n)]$$

即为所求. 但是也应同时指出除了某些特殊情形 (如正态分布的情形) 外, 还没有很好的实际可行方法求出函数 $\mathbf{E}[X|\sigma(X_1,\cdots,X_n)]$.

因此为了实际应用, 人们退一步将问题改成下面的形式: 设 $X,X_1,\cdots,X_n$ 是二阶矩存在的随机变量, 希望求 $X_1,\cdots,X_n$ 的一个线性函数 $\sum_{k=1}^n a_kX_k$, 使得对任何常数 $b_1,\cdots,b_n$, 有

$$\mathbf{E}|X-\sum_{k=1}^n a_kX_k|^2\le\mathbf{E}|X-\sum_{k=1}^n b_kX_k|^2 \tag{1}$$

这就是所谓的线性回归问题 (也就是最佳线性均方逼近问题, 有时也称为最小二乘原理) 当 $X_1,\cdots,X_n$ 为实随机变量时, 线性回归问题可以用古典分析方法解决. 为了与定理 3 的结论比较, 有时称满足 (1) 的 $\sum_{k=1}^n a_kX_k$ 为 X 在 $X_1,\cdots,X_n$ 下的广义条件期望, 且记作 $\hat{\mathbf{E}}[X|X_1,\cdots,X_n]$. 更一般地, 令 $\mathfrak{M}$ 为 $L^2(\Omega,\mathcal{F},\mathbf{P})$ 的一个线性子空间, $X\in L^2(\Omega,\mathcal{F},\mathbf{P})$ 在 $\mathfrak{M}$ 下的 **广义条件期望** 是指 $\mathfrak{M}$ 中满足条件

$$\mathbf{E}|X-\hat{\mathbf{E}}[X|\mathfrak{M}]|^2=\inf\{\mathbf{E}|X-f|^2:\ f\in\mathfrak{M}\}$$

的函数 $\hat{\mathbf{E}}[X|\mathfrak{M}]$. 广义条件期望与条件期望有很多类似的性质, 不再在此讨论了.

§5. 概率测度的收敛

在现代概率论中, 研究在一定意义下概率测度的收敛性以及 r.v. 序列的分布的收敛性是很有意义的. 例如, 它是描述一些随机现象随时间的演化达到某种平衡状态等物理概念的有力工具.

我们在定理 2.1.4 的前言中，曾提出过概率分布函数弱收敛的概念. 现在正式给出有关定义. 我们首先讨论 $(\mathbb{R},\mathcal{B})$ 上分布函数及概率测度的收敛概念.

1. 定义 设 $\{F_n\}, F$ 是 $\mathbb{R}$ 上一致有界分布函数序列 (不一定是概率分布函数), $C(F)$ 表示 F 的一切连续点集, 如果

$$F_n(x) \to F(x), \quad \forall\, x \in C(F),$$

则称 $\{F_n\}$**淡收敛** (vague convergence) 向 F, 记作

$$F_n \xrightarrow{v} F.$$

若还有 $F_n(\pm\infty) \to F(\pm\infty)$, 则称 $\{F_n\}$**弱收敛** (weak convergence) 向 F, 记作

$$F_n \Longrightarrow F \text{ 或 } F_n \xrightarrow{w} F.$$

设 $\{\mu_n\}$、μ 是 $(\mathbb{R}.\mathcal{B})$ 上一致有界测度序列, 若对一切 μ 的连续区间 $(a,b]$(即 $\mu(\{a,b\}) = 0$) 有

$$\mu_n\big((a,b]\big)\to \mu\big((a,b]\big),$$

则称 μ_n 淡收敛向 μ, 记作

$$\mu_n \xrightarrow{v} \mu.$$

若还有 $\mu_n(\mathbb{R}) \to \mu(\mathbb{R})$, 则称 $\{\mu_n\}$**弱收敛**(weak convergence) 向 μ, 记作

$$\mu_n \Longrightarrow \mu \text{ 或 } \mu_n \xrightarrow{w} \mu.$$

若 r.v. 序列 $\{X_n\}$ 相应的分布函数序列 $\{F_n\}$ 弱收敛向 r.v. X 的分布函数 F, 则称 $\{X_n\}$**依分布收敛** 向 X, 记作

$$X_n \Longrightarrow X \quad \text{或} X_n \xrightarrow{\mathcal{D}} X.$$

易证淡收敛的极限是唯一的 (对 r.v. 序列而言指分布唯一).

这里需要说明为什么我们不要求对每一 Borel 集 A, $\lim_{n\to\infty}\mu_n(A)=\mu(A)$, 下面的简单例子可以告诉我们理由：

例1 设 r.v. $X_n=1/n$, 显然 $X_n\to X, X=0$. 若 μ_n,μ 分别表示 X_n, X 的分布, 则对任何 $a<b, a\neq 0, b\neq 0$ 均有 $\mu_n\big((a,b)\big)\to\mu\big((a,b)\big)$, 但 $1=\mu_n\big(((0,1)\big)\neq\mu\big((0,1)\big)=0$.

再者，即便 μ_n 是概率测度，在某种意义下收敛，其极限也未必是概率测度. 请看下例.

例2 设 r.v. $X_n=n$, 显然 $X_n\to\infty$. 若 μ_n 表示 X_n 的分布，则显然对一切区间 $(a,b]$, 都有 $\mu_n\big((a,b]\big)\to 0$, 于是 $\mu_n\xrightarrow{v}\mu$, μ 是恒为零的测度.

由于淡收敛和弱收敛的差别只在于 $\pm\infty$ 处或 $\mathbb{R}$ 上的测度的极限的状况. 而淡收敛又有下面的定理 2 那样有用的结果, 所以人们在研究分布的弱收敛之前, 常常先研究淡收敛 (关于淡收敛读者可参考 [Ch1]).

2. 定理 (Helly 选择定理) 对任意给定的一致有界的分布函数序列必存在淡收敛的子序列.

证明 设 F_n 为 $\mathbb{R}$ 上一致有界的分布函数序列, $\sup\{|F_n(x)|: x\in\mathbb{R}, n\in\mathbb{N}\}=M<\infty$. 设 D 是 $\mathbb{R}$ 上可数稠密集, $\{r_k, k\geqslant 1\}$ 是它的一种排列, 数列 $\{F_n(r_1), n\geqslant 1\}$ 是有界的, 根据有界实数集的列紧性, 必存在所给序列的一个子列 $\{F_{1k}, k\geqslant 1\}$, 使得极限 $\lim_{k\to\infty}F_{1k}(r_1)=l_1$ 存在, 且 $|l_1|\leqslant M$. 其次, 数列 $\{F_{1k}(r_2), k\geqslant 1\}$ 是有界的, 故存在 $\{F_{1k}, k\geqslant 1\}$ 的子列 $\{F_{2k}, k\geqslant 1\}$, 使得 $\lim_{k\to\infty}F_{2k}(r_2)=l_2$ 存在, 且 $|l_2|\leqslant M$. 因为 $\{F_{2k}, k\geqslant 1\}$ 是 $\{F_{1k}, k\geqslant 1\}$ 的子列, 故它在 r_1 处也收敛于 l_1. 继续这样作下去, 我们得到

$F_{11}, F_{12}, \cdots, F_{1k}, \cdots$ 在 r_1 处收敛;

$F_{21}, F_{22}, \cdots, F_{2k}, \cdots$ 在 r_1, r_2 处收敛;

....

$F_{j1}, F_{j2}, \cdots, F_{jk}, \cdots$ 在 $r_1, r_2, \cdots, r_j$ 处收敛;

....

考虑对角线 $\{F_{kk},\ k \geqslant 1\}$, 易知它在每个 r_j 处收敛. 事实上, 对任何 j 来说, 除前面的 $j-1$ 项外, 序列 $\{F_{kk},\ k \geqslant 1\}$ 是 $\{F_{jk},\ k \geqslant 1\}$ 的一个子序列, 它在 r_j 处收敛, 故 $\lim_{k\to\infty} F_{kk}(r_j) = l_j$.

于是我们证明了存在一个子列 $\{n_k\}$ 及定义在 D 上的一个非降函数 G, 使得

$$\lim_{k\to\infty} F_{n_k}(r) = G(r), \quad \forall\ r \in D.$$

由 G 如下定义一个 $\mathbb{R}$ 上的函数 F:

$$F(x) = \inf_{x<r\in D} G(r), \quad \forall\, x \in \mathbb{R}.$$

显然, F 是非降的. 对每一 x 和 ε, 存在 $x < r \in D$, 使得 $G(r) < F(x) + \varepsilon$, 如果 $x \leqslant y < r$, 则 $F(y) \leqslant G(r) < F(x) + \varepsilon$, 因此 F 右连续.

如果 F 在 x 处连续, 选择 $y < x$ 使得 $F(x) - \varepsilon < F(y)$; 再取 r 和 s 使得 $y < r < x < s$ 和 $G(s) < F(x) + \varepsilon$ 成立. 从 $F(x) - \varepsilon < G(r) \leqslant G(s) < F(x) + \varepsilon$ 和 $F_n(r) \leqslant F_n(x) \leqslant F_n(s)$ 得出

$$\begin{aligned} F(x) - \varepsilon < G(r) = \lim_{k\to\infty} F_{n_k}(r) &\leqslant \varliminf_{k\to\infty} F_{n_k}(x) \\ &\leqslant \varlimsup_{k\to\infty} F_{n_k}(x) \leqslant \lim_{k\to\infty} F_{n_k}(s) \leqslant G(s) < F(x) + \varepsilon. \end{aligned}$$

令 $\varepsilon \to 0$, 即得 $\lim_{k\to\infty} F_{n_k}(x) = F(x)$. 这就证明了 $F_{n_k} \xrightarrow{v} F$. □

3. 推论 1) $(\mathbb{R}, \mathcal{B})$ 上任意一致有界测度序列有淡收敛子序列.

2) $\mu_n \xrightarrow{v} \mu$ 的充要条件是它的一切子序列有淡收敛向 μ 的子序列.

证明 1) 令

$$F_n(x)=\mu_n\big((-\infty,\ x]\big),\ \forall\ x\in\mathbb{R}.$$

则 $\sup\{F_n(x):\ x\in\mathbb{R},\ n\geqslant 1\}\leqslant \sup_n \mu_n(\mathbb{R})<\infty$. 由定理 2 知存在子序列 $F_{n_k}\xrightarrow{v}$ 某一 F. 取 $\mu=\mu_F$(与不降右连续函数 F 所对应的测度), 则对一切 μ 的连续区间 $I=(a,b]$, 必有 $a,b\in C(F)$, 因而

$$\begin{aligned}\lim_{k\to\infty}\mu_{n_k}\big((a,b]\big)&=\lim_{k\to\infty}\big(F_{n_k}(b)-F_{n_k}(a)\big)\\&=F(b)-F(a)=\mu\big((a,b]\big),\end{aligned}$$

即当 $k\to\infty$ 时, $\mu_{n_k}\xrightarrow{v}\mu$. 因而 1) 获证.

2) 条件的必要性是显然的, 往证条件的充分性. 若 $\{\mu_n\}$ 的一切子序列都有淡收敛向 μ 的子序列, 而 $\{\mu_n\}$ 不淡收敛向 μ, 则必存在一个 μ 连续区间 I, 使得 $\mu_n(I)\not\to\mu(I)$, 因而存在子列 $\{n_k\}$, 使 $\mu_{n_k}(I)\to h\neq\mu(I)$, 于是 μ_{n_k} 就不可能有淡收敛向 μ 的子序列. 这一矛盾就证明了条件的充分性. □

在 $\mathbb{R}$ 的连续函数类中记

$C=$ 连续函数类;

$C_b=$ 有界连续函数类;

$C_0=\{f\in C:\ 使得\ \lim\limits_{|x|\to\infty}f(x)=0\}$.

4. 定理 $\mu_n\xrightarrow{v}\mu$ 的充要条件是下述二条件之一成立:

1) 对一切 $f\in C_b$ 及 μ 的连续区间 I,

$$\int_I f(x)\ \mu_n(dx)\to\int_I f(x)\ \mu(dx);$$

2) 对一切 $f\in C_0$,

$$\int f(x)\ \mu_n(dx)\to\int f(x)\ \mu(dx).$$

证明 设 $\sup_n\{\mu_n(\mathbb{R})\} \leqslant M$, $\mu_n \xrightarrow{v} \mu$,

1) 若 $I=(a,b]$ 是 μ 的任一连续区间, $\sup_x|f(x)| \leqslant N$. f 在 $[a,b]$ 上一致连续, 因而任给 $\varepsilon>0$, 总存在 $\delta>0$, 使当 $x,x' \in [a,b]$ 且 $|x-x'|<\delta$ 时, $|f(x)-f(x')|<\varepsilon$. 再设 F 是 μ 的分布函数, F 的不连续点至多可数, 存在 $a=a_0<a_1<a_2<\cdots<a_m=b$, 对每一 k, $a_k-a_{k-1}<\delta$. 且 $a_k \in C(F)$, 从而 $I_k=(a_{k-1},a_k]$ 为 μ 的连续区间. 令

$$f^*(x)=\sum_{k=1}^m f(a_k)I_k(x),\ x\in[a,b],$$

则 f^* 是有界 Borel 可测函数, 于是

$$\begin{aligned}\left|\int_I f\,d\mu-\int_I f\,d\mu_n\right| &\leqslant \int_I|f-f^*|\,d\mu\\ &+\left|\int_I f^*\,d\mu-\int_I f^*\,d\mu_n\right|+\int_I|f-f^*|\,d\mu_n\\ &\leqslant \varepsilon M+\sum_{k=1}^m|f(a_k)||\mu(I_k)-\mu_n(I_k)|+\varepsilon M\\ &\leqslant 2\varepsilon M+N\sum_{k=1}^m|\mu(I_k)-\mu_n(I_k)|.\end{aligned}$$

由于每一 I_k 是 μ 的连续区间, 故当 $n\to\infty$ 时,

$$\overline{\lim_{n\to\infty}}\left|\int_I f\,d\mu-\int_I f\,d\mu_n\right|\leqslant 2\varepsilon M.$$

再由 ε 的任意性, 即知

$$\overline{\lim_{n\to\infty}}\left|\int_I f\,d\mu-\int_I f\,d\mu_n\right|=0.$$

2) 若 $f\in C_0$, 则 $f\in C_b$. 任给 $\varepsilon>0$, 存在 μ 连续区间 I, 使得当 $x\in I^c$ 时, $|f(x)|<\frac{\varepsilon}{3M}$, 因而

$$\left|\int f\,d\mu-\int f\,d\mu_n\right|$$

$$\leqslant \int_{I^c} |f|\, d\mu + \left| \int_I f\, d\mu - \int_I f\, d\mu_n \right| + \int_{I^c} |f|\, d\mu_n$$
$$\leqslant \frac{\varepsilon}{3M}\mu(I^c) + \left| \int_I f\, d\mu - \int_I f\, d\mu_n \right| + \frac{\varepsilon}{3M}\mu_n(I^c).$$

令 $n \to \infty$, 即得

$$\varlimsup_{n\to\infty} \left| \int f\, d\mu - \int f\, d\mu_n \right| \leqslant \frac{2\varepsilon}{3}.$$

由 ε 的任意性可得

$$\lim_{n\to\infty} \int f\, d\mu_n = \int f\, d\mu. \quad \square$$

淡收敛的概念可以扩展到 $(\mathbb{R}^k, \mathcal{B}^k)$ 上. 此时 μ–连续区间 $[a,b]$ 是指 $\mu([a,b] \setminus (a,b)) = 0$. 我们不在此仔细讨论了，读者可参看 [YWL] 第六章 §3.2.

下面我们讨论一般距离空间上概率测度的弱收敛 (读者可参考 [Bi2]).

5. 定义 设 $\{\mu, \mu_n, n \in \mathbb{N}\}$ 为距离空间 $(E, \mathcal{B})$ 中的概率测度序列，记 $C_b(E)$ 为 E 上有界连续函数集合，若

$$\int_E f(x)\, \mu_n(dx) \to \int_E f(x)\, \mu(dx), \forall\, f \in C_b(E),$$

则称 $\{\mu_n\}$**弱收敛**(weak convergence) 向 μ, 记作

$$\mu_n \Longrightarrow \mu \text{ 或 } \mu_n \xrightarrow{w} \mu.$$

易证定义 1 所定义的 $(\mathbb{R}, \mathcal{B})$ 上概率测度的弱收敛与本定义一致. 证明留作习题.

6. 引理 设 (E, ρ) 为一距离空间，若 F 是闭集，则对任何 $\varepsilon > 0$, 存在有界一致连续函数 $f \in C_b(E)$, 使得

(i) 对一切 $x \in F, f(x) = 1$;

(ii) 当 $\rho(x, F) > \varepsilon$ 时， $f(x) = 0$;

(iii) 对一切 $x \in E, 0 \leqslant f(x) \leqslant 1$.

函数 f 可以选成一致连续的.

证明 如下定义 $\mathbb{R}$ 上连续函数

$$\varphi(t) = \begin{cases} 1, & 若 \ t \leqslant 0, \\ 1-t, & 若 \ 0 \leqslant t \leqslant 1, \\ 0, & 若 \ 1 \leqslant t. \end{cases}$$

若令

$$f(x) = \varphi\Big(\frac{1}{\varepsilon}\rho(x, F)\Big),$$

则 f 具有所要求的性质，而且是一致连续的. □

7. 定理 设 μ 和 ν 是 $(E, \mathcal{B})$ 上的概率测度，若

$$\int_E f(x)\ \mu(dx) = \int_E f(x)\ \nu(dx),\ \forall\ f \in C_b(E),$$

则 $\mu = \nu$.

证明 只需证对一切闭集 F 有

$$\mu(F) = \nu(F).$$

因为闭集类为 $\pi-$ 系且生成 $\mathcal{B}$. 令

$$\Lambda := \{B \in \mathcal{B} : \mu(B) = \nu(B)\}.$$

显然 Λ 是 $\lambda-$ 系，因而定理获证.

对任意闭集 F, 令

$$f_n(x) = \varphi\big(n\rho(x, F)\big), \quad x \in E,\ n \in \mathbb{N},$$

则它是 E 上有界 (且一致) 连续函数，其中 φ 为引理 6 所定义，ρ 为 E 中距离. 由

$$\lim_{n\to\infty} f_n(x) \to I_F(x)$$

及控制收敛定理，即得

$$\mu(F)=\lim_{n\to\infty}\int f_n(x)\mu(dx)=\lim_{n\to\infty}\int f_n(x)\nu(dx)=\nu(F).\quad\Box$$

这一定理说明弱收敛的极限是唯一的.

8. 定理 设 $\{\mu,\ \mu_n,\ n\in\mathbb{N}\}$ 是 $(E,\mathcal{B})$ 上概率测度序列，则下述五个条件等价：

1) $\mu_n\Longrightarrow\mu$;

2) $\lim_{n\to\infty}\int f\,d\mu_n=\int f\,d\mu$, 对一切 $f\in U(E)$(E 上有界一致连续实函数类);

3) $\overline{\lim\limits_{n\to\infty}}\,\mu_n(F)\leqslant\mu(F)$, 对一切闭集 F;

4) $\underline{\lim\limits_{n\to\infty}}\,\mu_n(G)\geqslant\mu(G)$, 对一切开集 G;

5) $\lim\limits_{n\to\infty}\mu_n(A)\to\mu(A)$, 对一切 $A\in\mathcal{B}$ 且 $\mu(\partial A)=0$.

证明 1)$\Longrightarrow$2). 由 $U(E)\subset C_b(E)$ 即得；

2)$\Longrightarrow$3). 由引理 6 知，对任意闭集 F 及 $m\in\mathbb{N}$, 存在 $f_m\in U(E)$, 使得对一切 $x\in F, f_m(x)=1$, 而当 $\rho(x,F)>\dfrac{1}{m}$ 时, $f(x)=0$, 其中 ρ 为 E 中距离，且对一切 $x\in E, 0\leqslant f_m(x)\leqslant 1$. 易知 $f_m\to I_F$. 于是对一切 $m\in\mathbb{N}$,

$$\overline{\lim_{n\to\infty}}\,\mu_n(F)\leqslant\overline{\lim_{n\to\infty}}\int f_m\,d\mu_n.$$

由 2) 及 $f_m\in U(E)$ 知

$$\lim_{n\to\infty}\int f_m\,d\mu_n=\int f_m\,d\mu.$$

令 $m\to\infty$, 利用控制收敛定理即得 3) 成立.

3)$\Longleftrightarrow$4). 对任意开集 G, G^c 是闭集，由于 μ_n,μ 都是概率测度，因而 3) 与 4) 等价.

4)$\Longrightarrow$5). 对任意 $A \in \mathcal{B}, \mu(\partial A) = 0$, 有

$$A^{\circ} \subset A \subset \bar{A}$$

且

$$\mu(A^{\circ}) = \mu(\dot{A}) = \mu(\bar{A}),$$

其中集 A° 表示集 A 的内部，它是开集. 从而有

$$\mu(A) = \mu(A^{\circ}) \leqslant \varliminf_{n\to\infty} \mu_n(A^{\circ}) \leqslant \varliminf_{n\to\infty} \mu_n(A)$$
$$\leqslant \varlimsup_{n\to\infty} \mu_n(A) \leqslant \varlimsup_{n\to\infty} \mu_n(\bar{A}) \leqslant \mu(\bar{A}) = \mu(A).$$

此即

$$\lim_{n\to\infty} \mu_n(A) = \mu(A).$$

5)$\Longrightarrow$1). 对任意 $f \in C_b(E)$, 设 $a < f < b$, 由于 μ 有限，从而对任意正整数 k, 使得 $\mu(\{x : f(x) = c\}) > \frac{1}{k}$ 的实数只有有限个，从而使得 $\mu(\{x : f(x) = c\}) > 0$ 的 c 最多是可列个. 故对任意 $\varepsilon > 0$, 能适当地选取正整数 N 及实数 $a_j, j = 1, \cdots, N$, 使得 i)$a = a_0 < a_1 < \cdots < a_N = b$; ii) $a_j - a_{j-1} < \varepsilon,\ j = 1, \cdots, N$, 和 iii)$\mu(\{x : f(x) = a_j\}) = 0,\ j = 0, 1, \cdots, N$. 令

$$A_j = \{x :\ a_{j-1} \leqslant f(x) < a_j\}, \quad j = 1, \cdots, N.$$

易见每一 $A_j \in \mathcal{B}$ 且两两不交，且 $E = \bigcup_{j=1}^{N} A_j$. 此外，由

$$\partial A_j = \bar{A}_j - A_j^{\circ}$$
$$\subset \{x :\ f(x) = a_{j-1}\} \bigcup \{x :\ f(x) = a_j\}$$

得

$$\mu(\partial A_j) = 0, j = 1, \cdots, N,$$

即每一 A_j 是 μ-连续集. 令简单函数

$$f^* = \sum_{j=1}^{N} a_{j-1} I_{A_j},$$

则

$$|f(x)-f^*(x)|<\varepsilon, \quad \forall x\in E.$$

因此，

$$\begin{aligned}|\int f\,d\mu_n-\int f\,d\mu|\leqslant &|\int f\,d\mu_n-\int f^*\,d\mu_n|\\ &+|\int f^*\,d\mu_n-\int f^*\,d\mu|+|\int f^*\,d\mu-\int f\,d\mu|\\ &\leqslant \varepsilon\mu_n(E)+\sum_{j=1}^{N}|a_{j-1}||\mu_n(A_j)-\mu(A_j)|+\varepsilon\mu(E).\end{aligned}$$

因为 E 和 A_j, $j=1,\cdots,N$ 都是 μ–连续集，所以对上述不等式两边取上极限，即有

$$\overline{\lim_{n\to\infty}}|\int f\,d\mu_n-\int f\,d\mu|\leqslant 2\varepsilon\mu(E)=2\varepsilon.$$

由于 ε 的任意性，即得

$$\lim_{n\to\infty}|\int f\,d\mu_n-\int f\,d\mu|=0.$$

即得证 1). □

9. 定义 称 $(E,\mathcal{B})$ 上概率测度序列 $\{\mu_n\}$ 是 **相对紧** (relatively compact) 的，如果它的一切子序列都有弱收敛的子序列.

10. 定义 称 $(E,\mathcal{B})$ 上概率测度序列 $\{\mu_n\}$ 是 **胎紧**(tight) 的，如果对任何 $\varepsilon>0$, 存在紧集 K, 使得对一切 n, $\mu_n(K^c)<\varepsilon$.

11. 引理 若 E 为完全可分距离空间，则 $(E,\mathcal{B})$ 上的每一概率测度 μ 是胎紧的.

证明 因为 E 可分，所以对每一 n, 存在半径为 $1/n$ 的开球序列 $A_{n1},A_{n2},\cdots$ 覆盖 E, 任给 $\varepsilon>0$, 选择 i_n 使得 $\mu(\bigcup_{i\leqslant i_n}A_{ni})>1-\varepsilon/2^n$, 根据完全性假设，全有界集 $\bigcap_{n\geqslant 1}\bigcup_{i\leqslant i_n}A_{ni}$ 的闭包 K 为紧集 (参看定理 2.4.10). 显然 $\mu(K)>1-\varepsilon$. 因而 μ 是胎紧的.

□

12. 定理 (Прохоров 定理) 若 E 是完全可分的，则 $\{\mu_n\}$ 相对紧等价于胎紧.

这一定理的证明比较复杂，有兴趣的读者可参看 [Bi] pp.37–40 或 [LLL] pp.192–196. .

13. 定理 $\mu_n \Longrightarrow \mu$ 的充要条件是 $\{\mu_n\}$ 的任意子序列 $\{\mu_{n'}\}$ 有子序列 $\mu_{n''} \Longrightarrow \mu$.

证明 将推论 3.2) 的证明中的 μ 连续区间代之以 μ–连续集，即得本定理的证明.

14. 定理 $(\mathbb{R}^k, \mathcal{B}^k)$ 上概率测度序列 μ_n 弱收敛的充分必要条件是 $\mu_n, n \geqslant 1$ 胎紧且淡收敛.

证明留作习题.

习题

1. 证明淡收敛的极限是唯一的.

2. 证明定义 1 所定义的 $(\mathbb{R}, \mathcal{B})$ 上概率测度的弱收敛与定义 5 一致.

3. 证明定理 14.

4. 证明 $\mathbb{R}$ 上概率分布函数族 $\{F_\alpha\}$ 所对应的概率测度族相对紧的充分必要条件是当 $x \to -\infty$ 和 $x \to \infty$ 时，$\{F_\alpha\}$ 对 α 一致收敛.

5. 设与随机变量族 $\{X_\alpha\}$ 相应的概率分布族为 $\{\mu_\alpha\}$. 如果对某个实数 $r > 0$,$\mathbf{E}\{|X_\alpha|^r\}$ 对 α 有界，则 μ_α 相对紧.

§6 几个收敛之间的关系的注记

由前面几节讨论知概率测度空间 (有限测度空间) 上几种收敛的关系可以概括如下:

$$X_n \xrightarrow{\text{a.s.}} X \implies X_n \xrightarrow{\mathbf{P}} X \Longleftarrow X_n \xrightarrow{\mathcal{D}} X$$

$$\Downarrow$$

$$X_{n_k} \xrightarrow{\text{a.s.}} X \text{ 且} \sum_{k=1}^{\infty} \mathbf{P}\left(|X_{n_k} - X| \geq \frac{1}{2^k}\right) < \infty$$

$$\Uparrow$$

$$X_n \xrightarrow{r} X \implies X_n \xrightarrow{r'} X, \quad 0 < r' < r$$

"$\xrightarrow{\mathcal{D}}$" 不能推出 "$\xrightarrow{\mathbf{P}}$" 是显然的. 现在举几个反例说明 "$\xrightarrow{\mathbf{P}}$" 不能推出 "$\xrightarrow{r}$", 也不能推出 "$\xrightarrow{\text{a.e.}}$", "$\xrightarrow{\text{a.e.}}$" 不能推出 "$\xrightarrow{r}$"; 而 "$\xrightarrow{r}$" 也不能推出 "$\xrightarrow{\text{a.e.}}$".

例1 设 $(\Omega, \mathcal{F}, \mathbf{P}) = ((0,1], \mathcal{B}(0,1], \lambda)$.

$$\varphi_{k,j} := I_{\left(\frac{j-1}{k}, \frac{j}{k}\right]}, \quad k \in \mathbb{N}, \quad j = 1, \cdots, k$$

按字典式的次序 (先按 k 增加的顺序, 再对每个 k 按 j 增加排序), 将它们排成 X_n, 于是如果 $X_n = \varphi_{k_n, j_n}$, 则

$$\lim_{n \to \infty} \mathbf{E} X_n^p = \lim_{n \to \infty} \frac{1}{k_n} = 0.$$

即 $X_n \xrightarrow{\mathbf{P}} 0$, 但对一切 $\omega \in \Omega$ $k \in \mathbb{N}$, 存在 $j \in \mathbb{N} \cap [1,k]$ 使 $\varphi_{kj}(\omega) = 1$, 所以有无穷多个 n 使 $X_n(\omega) = 1$, 同理有无穷多个 n 使 $X_n(\omega) = 0$. 故对一切 $\omega \in (0,1]$, $\{X_n(\omega)\}$ 不可能收敛.

即 $X_n \xrightarrow{\mathbf{P}} 0$ 不能推出 $X_n \xrightarrow{\text{a.e.}} 0$.

例2 . 对一切 $k \in \mathbb{N}$, $j = 1, 2, \cdots, k$, $p > 0$, 令

$$\psi_{kj} = k^{1/p} \varphi_{kj},$$

其中 $\varphi_{k,j}$ 如例 1 定义，再按字典式将 ψ_{kj} 排序，记成 $\{Y_n\}$. 则 $\mathbf{P}(|Y_n|>0)=\mathbf{P}(X_n>0)=1/k_n\to 0$, 即 $Y_n\xrightarrow{\mathbf{P}}0$. 但

$$\mathbf{E}Y_n^p=k_n\mathbf{E}X_n^p=k\cdot\frac{1}{k_n}\equiv 1,\quad \forall\, n\in\mathbb{N}.$$

故 $Y_n\xrightarrow{\mathbf{P}}0$, 但 $Y_n\overset{p}{\not\to}0$.

例3 . 在例 1 的概率空间上，令

$$X_n(\omega):=n^{1/p}I_{(0,\frac{1}{n})}.$$

则 $X_n\xrightarrow{\text{a.e.}}0$, 但 $\mathbf{E}X_n^p=n\cdot\mathbf{E}I_{(0,\frac{1}{n})}=n\cdot\frac{1}{n}=1\not\to 0$, 即 $X_n\overset{p}{\not\to}0$.

在随机过程中有更多自然产生的例子.

第九章 大数定律、随机级数

本章将应用第八章的各种收敛概念讨论 "大数定律"— 这是概率论中的一个重要部分，但是我们这里主要讨论 r.v. 独立或不相关的情况.

§1. 简单的极限定理及其应用

1. **定义** 设 $\{X_n: n \in \mathbb{N}\}$ 是一 r.v. 序列，令

$$S_n := \sum_{k=1}^{n} X_k, \qquad n \in \mathbb{N},$$

在古典意义下，若 $\mathbf{E}(S_n)$ 有限且

$$\frac{S_n - E(S_n)}{n} \xrightarrow{\mathbf{P}} 0,$$

则称 $\{X_n\}$ 服从 **弱大数律**;

$$\frac{S_n - E(S_n)}{n} \xrightarrow{\text{a.s.}} 0,$$

则称 $\{X_n\}$ 服从 **强大数律**. 在现代意义下，若存在 $\{a_n\} \subset \mathbb{R}$, $\{b_n\} \subset \mathbb{R}_+$, 且 $b_n \to \infty$ 使

$$\frac{S_n - a_n}{b_n} \to 0 \quad 依概率\mathbf{P}(或\ \text{a.s.}),$$

则称 $\{X_n\}$ 服从弱 (或强) 大数律.

2. 由定理 8.3.2 立即得到

$$E(S_n^2) = o(n^2) \implies \frac{S_n}{n} \xrightarrow{\mathbf{P}} 0, \tag{1}$$

因而由定理 8.2.4 知有一子列 $\{n_k\}\subset\mathbb{N}$ 使

$$\frac{S_{n_k}}{n_k}\xrightarrow{\text{a.s.}}0.$$

下面我们讨论一下什么条件下 $E(S_n^2)=o(n^2)$ 成立：容易算出

$$\mathbf{E}S_n^2=\sum_{k=1}^{n}\mathbf{E}(X_n^2)+2\sum_{1\leqslant j<k\leqslant n}\mathbf{E}(X_jX_k), \tag{2}$$

这里右边有 n^2 项，如果不加条件，一般只能得出 $\mathbf{E}S_n^2=O(n^2)$. 要想得到 (1) 中的前提, 最自然的想法是引进下列的概念 (它在现实中也是有意义的).

定义 设 X,Y 为 r.v. , $\mathbf{E}|X|^2<\infty$, $\mathbf{E}|Y|^2<\infty$ 且

$$\mathbf{E}(XY)=\mathbf{E}X\cdot\mathbf{E}Y, \tag{3}$$

则称 X,Y**不相关**，若 (3) 代以

$$\mathbf{E}(XY)=0, \tag{4}$$

则称 X,Y**正交**.

注意： (3) 等价于 $\operatorname{cov}(X,Y)=\mathbf{E}(X-\mathbf{E}X)(Y-\mathbf{E}Y)=0$. 即 X,Y 的相关矩为 0. 当 $\mathbf{E}(X)=\mathbf{E}(Y)=0$ 时， (3) 化归 (4). 而二阶矩有限的条件是需要的，因为它保证 $\mathbf{E}(XY)$, $\mathbf{E}X$, $\mathbf{E}Y$ 的有限. 还应注意

$$X,Y\text{独立}\implies X,Y\text{不相关}. \tag{5}$$

若 $\{X_n:n\in\mathbb{N}\}$ 为两两不相关 r.v. 列，则 $\{X_n-\mathbf{E}X_n:n\in\mathbb{N}\}$ 为正交列，于是由 (2) 的计算得到

$$\sigma^2(S_n)=\sum_{k=1}^{n}\sigma^2(X_n). \tag{6}$$

若 $\sup_n\sigma^2(X_n)<\infty$, 则得

$$\sigma^2(S_n)=O(n)=o(n^2)$$

因而由 (1) 并经直接计算即得

3. 定理 若 $\{X_n:\ n\in\mathbb{N}\}$ 两两不相关，且 $\sup_n \sigma^2(X_n)<\infty$. 则

$$\frac{S_n-\mathbf{E}S_n}{n}\xrightarrow{2}0.$$

因而

$$\frac{S_n-\mathbf{E}S_n}{n}\xrightarrow{\mathbf{P}}0.$$

事实上，我们可以将上述结果增强为

4. 定理 在定理 3 的相同假设下，

$$\frac{S_n-\mathbf{E}S_n}{n}\xrightarrow{\text{a.s.}}0.$$

证明 不失一般性，可设对一切 $j\in\mathbb{N}$, $\mathbf{E}(X_j)=0$. 于是，X_n, $n\in\mathbb{N}$, 正交， $\sup_n \mathbf{E}X_n^2=:M<\infty$. 而由 (2.4) 即得

$$\mathbf{E}(S_n^2)\leqslant Mn.$$

再由 чебышев 不等式知：对一切 $\varepsilon>0$,

$$\mathbf{P}(|S_n|\geqslant n\varepsilon)\leqslant\frac{Mn}{n^2\varepsilon^2}=\frac{M}{n\varepsilon^2},\tag{1}$$

如果我们对子列 $\{n^2\}$ 求和即得

$$\sum_{n=1}^{\infty}\mathbf{P}(|S_{n^2}|\geqslant n^2\varepsilon)\leqslant\sum_{n=1}^{\infty}\frac{M}{n^2\varepsilon^2}<\infty.$$

则由 Borel -Cantelli 引理知：

$$\mathbf{P}(|S_{n^2}|\geqslant n^2\varepsilon,\quad \text{i.o.})=0.$$

故对一切 $\varepsilon>0$ 有

$$\mathbf{P}\Big(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\Big\{\Big|\frac{S_{n^2}}{n^2}\Big|\geqslant\varepsilon\Big\}\Big)=0,$$

因而由定理 8.1.3(i) 知：

$$\frac{S_{n^2}}{n^2}\xrightarrow{\text{a.s.}}0.\tag{2}$$

现在我们将这个结果扩张到整个序列 S_n/n, $n \in \mathbb{N}$.

显然对一切 $k \in \mathbb{N}$, 存在 $n \in \mathbb{N}$, 使 $n^2 \leqslant k < (n+1)^2$, 而

$$|S_k| \leqslant |S_{n^2}| + |S_k - S_{n^2}| \leqslant |S_{n^2}| + D_n, \tag{3}$$

其中

$$D_n := \max_{n^2 \leqslant k < (n-1)^2} |S_k - S_{n^2}| = \max_{n^2 < k < (n+1)^2} |S_k - S_{n^2}|.$$

于是若能证明

$$\frac{D_n}{n^2} \xrightarrow{\text{a.s.}} 0, \tag{4}$$

则由 (3) 即得

$$\frac{|S_k|}{k} \leqslant \frac{|S_{n^2}| + D_n}{n^2}, \qquad n^2 \leqslant k < (n+1)^2,$$

结合 (2), (4) 即得定理的结论.

由 D_n 的定义知

$$D_n^2 = \max_{n^2 < k < (n+1)^2} |S_k - S_{n^2}|^2 \leqslant \sum_{k=n^2+1}^{(n-1)^2-1} |S_k - S_{n^2}|^2.$$

于是

$$\mathbf{E}(D_n^2) \leqslant \sum_{k=n^2+1}^{(n+1)^2-1} \sum_{j=n^2+1}^{k} \sigma^2(X_j) \leqslant 4n^2 M$$

因而由 чебышев 不等式得

$$\mathbf{P}(D_n \geqslant n^2\varepsilon) \leqslant \frac{\mathbf{E}(D_n^2)}{n^4\varepsilon^2} \leqslant \frac{4M}{n^2\varepsilon^2}$$

应用由 (1) 推出 (2) 的同样方法即得 (4). □

下面我们介绍定理 3.4 的一些应用.

5. 定理 (Borel) 对一切 $\omega \in [0,1)$ 将其展成十进小数

$$\omega = 0.\xi_1\xi_2\cdots, \quad \xi_k(\omega) \in \{0,1,2,\cdots,9\}$$

当 $\omega = m/10^n (n \in \mathbb{N},\ m \in [1, 10^n) \cap \mathbb{N})$ 时上述表示式取有限形式 (即某一位后都是数码 0). 则 ξ_k, $k \in \mathbb{N}$, 是 ω 的 Borel 可测函数, 令

$$\nu_\ell^{(n)}(\omega) := |\{k :\ \xi_k(\omega) = \ell,\ k = 1, 2, \cdots, n\}|,$$

其中 $|A|$ 表示集合 A 中元素的个数, 则

$$\lambda(\{\omega \in [0,1) :\ \lim_{n\to\infty} \frac{\nu_\ell^{(n)}(\omega)}{n} = \frac{1}{10},\quad \ell = 0, 1, \cdots, 9\}) = 1,$$

其中 λ 为 $\mathcal{B}[0,1)$ 上的 Lebesgue 测度.

注 这条定理就是有名的 **Borel 强大数律**, 它是解释概率概念的一个最古典的基本定理, 后来 Хинчин 用连分数证明了概率论中的第一个 **重对数律**, 这些对概率论的极限定理 (独立情况) 有深远的影响.

证明 (大意) 考虑概率空间 $(\Omega, \mathcal{F}, \mathbf{P}) := ([0,1), \mathcal{B}[0,1), \lambda)$, 则可证上述的 ξ_n, $n \in \mathbb{N}$, 是独立且同分布的 r.v. (iid), 其分布为

$$\mathbf{P}(\xi_n = \ell) = \frac{1}{10},\quad \ell = 0, 1, \cdots, 9,\quad n \in \mathbb{N},$$

进而, 对给定的 $\ell \in \{0, 1, \cdots, 9\}$, 定义

$$X_n := I_{\{\xi_n = \ell\}}.$$

则 X_n, $n \in \mathbb{N}$, 也是 iid, 因而满足定理 3 的条件 ($\mathbf{E}X_n = 1/10$, $\mathbf{E}X_n^2 = 1/10$), 于是 $\nu_\ell^{(n)}(\omega) = \sum_{k=1}^n X_k(\omega)$, 由定理 4

$$\frac{\nu_\ell^{(n)}}{n} = \frac{1}{n}\sum_{k=1}^n X_k \to \frac{1}{10} \text{ a.s.}. \qquad \square$$

6. Monte Carlo 积分 设 $f \in L^2([0,1]), \lambda)$, ξ_n, $n \in \mathbb{N}$, iid 且 $\xi_n \sim U[0,1]$ (此处及以后用 $U(I)$ 表示区间 I 上的均匀分布), 则

$$I_n := \frac{1}{n}\sum_{k=1}^n f(\xi_k) \xrightarrow[\text{a.s.}]{\mathbf{P}} \int_0^1 f(x)\,dx =: I,$$

且

$$\mathbf{P}(|I_n - I| > an^{-1/2}) \leqslant \frac{1}{a^2}\int_0^1 \left(f - \int_0^1 f dx\right)^2 dx. \quad \square$$

7. 例 高维立方体几乎是一球面.

设 $X_n, n \in \mathbb{N}$, 为 *iid* 序列, 且 $X_n \sim U(-1,1)$, 则 $\sigma^2(X_n) = \mathbf{E}X_n^2 = 1/3$, 所以由定理 4 知

$$\frac{1}{n}\sum_{k=1}^{n} X_k^2 \xrightarrow[\mathbf{P}]{\text{a.s.}} \frac{1}{3} \qquad (n \to \infty).$$

这就是说, 当 n 充分大时, $\mathbb{R}^n$ 的立方体 $(-1,1)^n$ 中的点几乎都在 $\mathbb{R}^n$ 的球面 $S(0,\sqrt{n/3})$ 附近.

或者说, 对一切 $\varepsilon > 0$,

$$\frac{1}{2^n}\lambda\left(\left\{x \in \mathbb{R}^n:\ \left(\frac{1}{3} - \varepsilon\right)n \leqslant |x|^2 \leqslant \left(\frac{1}{3} + \varepsilon\right)n\right\} \cap (-1,1)^n\right\}\right) \to 1.$$

8. 定理 (Weierstrass 逼近定理的 Bernstein 形式) 设 $f \in C([0,1])$, 令

$$B_n(x) := \sum_{m=0}^{n} \binom{n}{m} x^m (1-x)^{n-m} f\left(\frac{m}{n}\right),$$

$$x \in [0,1], \quad n \in \mathbb{N}. \qquad (n\text{阶 } \textbf{Bernstein 多项式})$$

则

$$\sup_{0 \leqslant x \leqslant 1} |B_n(x) - f(x)| \to 0.$$

即 f 可用它的 Bernstein 多项式一致逼近.

证明 对一切 $x \in [0,1]$, 存在 *iid* 序列 $X_n, n \in \mathbb{N}$, 使

$$\mathbf{P}(X_n = 1) = x, \qquad \mathbf{P}(X_n = 0) = 1 - x, \qquad n \in \mathbb{N},$$

因而 $\mathbf{E}X_n = x$, $\sigma^2(X_n) = x(1-x)$. 令 $S_n := \sum_{k=1}^{n} X_k$, 则

$$\mathbf{P}(S_n = m) = \binom{n}{m} x^m (1-x)^{n-m}, \quad m = 0, 1, \cdots, n,$$

因而 $\mathbf{E}f(S_n/n)=B_n(x)$. 由于 f 在 $[0,1]$ 上一致连续, 所以对一切 $\varepsilon>0$, 存在 $\delta>0$ 使当 $|x-y|<\delta$ 时, $|f(x)-f(y)|<\varepsilon$. 于是

$$
\begin{aligned}
|B_n(x)-f(x)| &= \left|\mathbf{E}f\left(\frac{S_n}{n}\right)-f(x)\right| \leqslant \int\left|f\left(\frac{S_n}{n}\right)-f(x)\right|d\mathbf{P}\\
&= \int_{\left\{\left|\frac{S_n}{n}-x\right|<\delta\right\}}\left|f\left(\frac{S_n}{n}\right)-f(x)\right|d\mathbf{P}\\
&\quad+\int_{\left\{\left|\frac{S_n}{n}-x\right|\geqslant\delta\right\}}\left|f\left(\frac{S_n}{n}\right)-f(x)\right|d\mathbf{P}\\
&\leqslant \varepsilon+2\sup_{y\in[0,1]}|f(y)|\mathbf{P}\left(\left|\frac{S_n}{n}-x\right|\geqslant\delta\right).
\end{aligned}
$$

再由 чебышев 不等式知

$$
\begin{aligned}
\mathbf{P}\left(\left|\frac{S_n}{n}-x\right|\geqslant\delta\right) &\leqslant \frac{1}{\delta^2}\sigma^2\left(\frac{S_n}{n}\right)\\
&= \frac{\sigma^2(X_1)}{n\delta^2}=\frac{x(1-x)}{n\delta^2}\leqslant\frac{1}{4\delta^2 n}.
\end{aligned}
$$

故由上两式即得

$$
\sup_{x\in[0,1]}|B_n(x)-f(x)|\leqslant\varepsilon+\frac{\sup_{x\in[0,1]}|f(x)|}{2\delta^2 n}. \qquad \square
$$

9. 定理 (Laplace 变换的逆转公式) 设 μ 为 $[0,\infty)$ 上的概率, 并令

$$
\varphi(\theta):=\int_0^\infty e^{-\theta y}\mu(dy),
$$

其中 $\int_0^\infty:=\int_{[0,\infty)}$. 则对一切 $x\in[0,\infty)$, $\mu(\{x\})=0$, 有

$$
\lim_{n\to\infty}\sum_{k=0}^{[nx]}\frac{(-1)^k}{k!}n^k\varphi^{(k)}(n)=\mu[0,x].
$$

即 μ 可由 φ 表出.

证明 对一切 $\theta>0$, 可以对 φ 求 k 阶导数

$$
\varphi^{(k)}(\theta)=\int_0^\infty(-y)^k e^{-\theta y}\mu(dy).
$$

取 $\theta=n$, $x>0$, 则

$$\sum_{k=0}^{[nx]}\frac{(-1)^k}{k!}n^k\varphi^{(k)}(n)=\int_0^\infty\sum_{k=0}^{[nx]}e^{-ny}\frac{(ny)^k}{k!}\mu(dy).$$

由于 $e^{-ny}(ny)^k/k!$ 是以 ny 为参数的 Poisson r.v. 等于 k 的概率, 想到取 $X_1,\cdots,X_n,\cdots$ *iid*. $X_n\sim\mathcal{P}(y)$(表示以 y 为参数的 Poisson 分布), 则由此假设及定理 3 得

$$\begin{aligned}\sum_{k=0}^{[nx]}e^{-ny}\frac{(ny)^k}{k!}&=\mathbf{P}(X_1+\cdots+X_n\leqslant nx)\\&=\mathbf{P}\Big(\frac{1}{n}\sum_{k=1}^{n}(X_k-\mathbf{E}X_k)\leqslant x-y\Big)\\&\longrightarrow\begin{cases}0, & x-y<0,\\1, & x-y>0,\end{cases}\qquad n\to\infty.\end{aligned}$$

故当 $\mu(\{x\})=0$ 时, 即得

$$\lim_{n\to\infty}\sum_{k=0}^{[nx]}\frac{(-1)^k}{k!}n^k\varphi^{(k)}(n)=\mu[0,x].\quad\Box$$

习题

1. 设 $(\Omega,\mathcal{F},\mathbf{P}_\theta)$, $\theta\in\Theta$, 为一族概率空间, Θ 为一有限或无限区间, X_n, $n\in\mathbb{N}$, 为 r.v. 列, 且

$$\mathbf{E}_\theta X_n=\theta,\qquad\sigma_n^2(\theta):=\mathbf{E}_\theta(X_n-\theta)^2=\sigma_\theta^2(X_n).$$

设 $u\in C_b(\Theta)$ (Θ 上的一切有界连续函数组成的集), 且对一切 $\theta\in\Theta$,$\sigma_n^2(\theta)\to0(n\to\infty)$, 试证:

$$\mathbf{E}_\theta u(X_n)\to u(\theta)\ (n\to\infty),\qquad\theta\in\Theta.$$

而且在 $\sigma_n^2(\theta)\to0$ 一致成立的每个闭区域上, 上述收敛是一致的.

试应用上述结论于下列各种情形, 并得出相应结论:

1) 设 $\mathbf{P}_\theta\Big(X_n=\dfrac{k}{n}\Big)=\dbinom{n}{k}\theta^k(1-\theta)^{n-k}$, $k=0,1,2,\cdots,n$;

2) 设 $\mathbf{P}_\theta\left(X_n = \frac{k}{n}\right) = e^{-n\theta}\frac{(n\theta)^k}{k!}$, $k = 0, 1, 2, \cdots, n$;

3) 设 X_n 在 $\mathbf{P}_\theta$ 之下服从参数为 n 和 n/θ 的 Γ 分布, 即

$$\mathbf{P}_\theta(X_n \leqslant x) = \begin{cases} 0, & x \leqslant 0, \\ \frac{1}{(n-1)!}\int_0^x \left(\frac{n}{\theta}\right)\left(\frac{nt}{\theta}\right)^{n-1} e^{-nt/\theta}dt, & x > 0. \end{cases}$$

对于 2), 3) 两种情形, n 为正实数时亦有相应结论, 不过在 3) 中 $(n-1)!$ 应用 $\Gamma(n)$ 代替.

2. 设 $f \in C[0,1]$, $0 \leqslant f \leqslant 1$, ξ_n, $n \in \mathbb{N}$, 为独立 r.v. , 且 $\xi_n \sim U[0,1]$, 对一切 $\varepsilon > 0$, 令 $\eta_\varepsilon(n) := I_{\{\xi_n \leqslant f(\varepsilon n)\}}$, $n \in [0, 1/\varepsilon]$, 试证:

$$\varepsilon \sum_{0 \leqslant \varepsilon n \leqslant 1} \eta_\varepsilon(n) \xrightarrow{\mathbf{P}} \int_0^1 f(r)dr \ (n \to \infty).$$

3. f 的假设同上题, 设 $X_1, \cdots, X_n, \cdots$ 为 *iid* , 且 $X_n \sim U[0,1]^2$, 设 $X_n = (X_{n1}, X_{n2})$, 则

$$\frac{1}{n}\sum_{k=1}^n I_{\{X_{n2} \leqslant f(X_{n1})\}} \xrightarrow[\text{a.s.}]{\mathbf{P}} \int_0^1 f(r)dr \ (n \to \infty).$$

4. f, ξ_n, $\eta_\varepsilon(n)$ 的假设与习题 2 同, 设 $\varphi \in C[0,1]$, 则

$$\varepsilon \sum_{0 \leqslant \varepsilon n \leqslant 1} \varphi(\varepsilon n)\eta_\varepsilon(n) \xrightarrow{\mathbf{P}} \int_0^1 \varphi(r)f(r)dr \ (n \to \infty).$$

5. 将定理 8.9 推广到 $f \in C[0,1]^d$ 的情形.

§2. 弱大数定律

定理 1.3 中的结论仅与 X_n 的期望有关, 但条件却涉及到二阶矩, 于是我们可以问: 这个假设是否可以去掉? 这样就有了 Хинчин 弱大数定律, 他证明这条定理的同时, 引导出等价序列的方法.

1. 定义 设 $\{X_n:\ n\in\mathbb{N}\}$, $\{Y_n:\ n\in\mathbb{N}\}$ 是两个随机变量序列，若

$$\mathbf{P}(X_n\neq Y_n,\ \text{i.o.})=0,$$

则称它们 **尾等价**.

由 Borel-Cantelli 定理立知，若

$$\sum_{n=1}^{\infty}\mathbf{P}(X_n\neq Y_n)<\infty,$$

则 $\{X_n:\ n\in\mathbb{N}\}$ 与 $\{Y_n:\ n\in\mathbb{N}\}$ 尾等价，由此易知：对于一个 r.v. 列 $\{X_n:\ n\in\mathbb{N}\}$, 常常可以通过"截断"的手法得到一个与之尾等价的 r.v. 列.

对于给定的 r.v. X 及 $c\in(0,\infty)$, 称

$$X^c:=XI_{\{|X|\leqslant c\}}=\begin{cases}X, & 当|X|\leqslant c\\ 0, & 当|X|>c\end{cases}$$

为 X 在 c 处的 **截断**. 于是对于任意一个有限值 r.v. 列 $\{X_n:\ n\in\mathbb{N}\}$, 存在 $c_n\in(0,\infty)$ 使

$$\mathbf{P}(|X_n|>c_n)<\frac{1}{2^n},$$

因而

$$\sum_{n=1}^{\infty}\mathbf{P}(X_n\neq X_n^{c_n})=\sum_{n=1}^{\infty}\mathbf{P}(|X_n|>c_n)<1.$$

于是 $\{X_n:\ n\in\mathbb{N}\}$ 与 $\{X_n^{c_n}:\ n\in\mathbb{N}\}$ 尾等价.

2. 引理 设 $\{X_n:\ n\in\mathbb{N}\}$, $\{Y_n:\ n\in\mathbb{N}\}$ 为尾等价的两个 r.v. 列，则

1) $\sum_{n=1}^{\infty}(X_n-Y_n)$ a.s. 收敛；

2) 对一切 $\{a_n\}\subset\mathbb{R}_+$, $a_n\to\infty$,

$$\frac{1}{a_n}\sum_{k=1}^{n}(X_k-Y_k)\to 0,\ \text{a.s.}$$

因而也依概率收敛于 0;

3) 以概率 1, $\sum X_n$ 与 $\sum Y_n$、$(1/a_n)\sum X_n$ 与 $(1/a_n)\sum Y_n$ 以同样的方式收敛或发散 (**收敛等价**);

4) 若 $(1/a_n)\sum_{k=1}^n X_k \xrightarrow{\mathbf{P}} X$, 则 $(1/a_n)\sum_{k=1}^n Y_k \xrightarrow{\mathbf{P}} X$.

证明 由尾等价定义知, 存在 $N \in \mathcal{F}$, $\mathbf{P}(N) = 0$, 使对一切 $\omega \in N^c$, 不可能有无穷个 $n \in \mathbb{N}$ 使 $X_n(\omega) \neq Y_n(\omega)$ 成立, 即存在 $n_0(\omega)$ 使

$$\forall\ n \geqslant n_0(\omega), \qquad X_n(\omega) = Y_n(\omega). \tag{1}$$

由此即知 1), 2), 3), 4) 成立.

3. 推论 若 $\sum_{n=1}^\infty \mathbf{P}(X_n \neq Y_n) < \infty$, 则引理 2 中的 1)–4) 都成立.

4. 定理 设 X_n, $n \in \mathbb{N}$, 为 *iid* 序列, $S_n := \sum_{k=1}^n X_k$, 则 $\{X_n : n \in \mathbb{N}\}$ 满足弱大数律 (即存在 $\{a_n\} \subset \mathbb{R}$ 使 $S_n/n - a_n \to 0$) 的充要条件是

$$\lim_{x\to\infty} x\mathbf{P}(|X_1| > x) = 0. \tag{1}$$

在条件成立时, $a_n := \mathbf{E}(X_1^n)$. (注意: 本定理的这一部分不要求期望 $\mathbf{E}X_1$ 存在.)

特别 (Хинчин 定理), 若 $\mathbf{E}|X_1| < \infty$, 则

$$\frac{S_n}{n} \xrightarrow{\mathbf{P}} \mathbf{E}X_1. \tag{2}$$

证明 充分性. 对给定的 n, 令

$$X_{nk} := X_k^n = X_k I_{\{|X_k| \leqslant n\}}, \quad k = 1, 2, \cdots, n,$$

$$S_n' := X_{n1} + \cdots + X_{nn}.$$

则 $X_{n1}, \cdots, X_{nn}$ 独立, $\mathbf{E}X_{nk} = a_n$ 且对任何给定的 $\varepsilon > 0$,

$$\mathbf{P}(|S_n/n - a_n| > \varepsilon) \leqslant \mathbf{P}(|S_n'/n - a_n| > \varepsilon) + \mathbf{P}(S_n \neq S_n').$$

右边第二项

$$
(3)\quad \begin{aligned} \mathbf{P}(S_n \neq S'_n) &\leqslant \mathbf{P}(\exists\, k \text{ 使 } X_k \neq X_{nk}) \\ &\leqslant \sum_{k=1}^{n} \mathbf{P}(X_k \neq X_{nk}) = n\mathbf{P}(|X_1| > n) \to 0. \quad (n \to \infty). \end{aligned}
$$

右边第一项由 чебышев 不等式知

$$
(4)\quad \mathbf{P}(|S'_n/n - a_n| > \varepsilon) \leqslant \varepsilon^{-2}\sigma^2\Big(\frac{S'_n}{n}\Big) = \frac{\sigma^2(X_{n1})}{n\varepsilon^2} \leqslant \frac{\mathbf{E}X_{n1}^2}{n\varepsilon^2},
$$

而由引理 6.2.11 知

$$
\begin{aligned} \mathbf{E}|X_{n1}|^2 &= \int_0^\infty x^2\, dF_{|X_{n1}|}(x) = 2\int_0^\infty y(1 - F_{|X_{n1}|}(y))\, dy \\ &= 2\int_0^\infty y\mathbf{P}(|X_{n1}| > y)\, dy \leqslant 2\int_0^n y\mathbf{P}(|X_1| > y)\, dy. \end{aligned}
$$

最后一个不等式是因为：当 $y \geqslant n$ 时，$\mathbf{P}(|X_{n1}| > y) = \mathbf{P}(0 > y) = 0$; 当 $y \leqslant n$ 时，

$$
\begin{aligned} \mathbf{P}(|X_{n1}| > y) &= \mathbf{P}(|X_{n1}| > n) + \mathbf{P}(n \geqslant |X_{n1}| > y) \\ &= \mathbf{P}(|X_1| \in (y, n]) \leqslant \mathbf{P}(|X_1| > y) \end{aligned}
$$

由假设知对任何 $\delta > 0$, 存在 $y_0 > 0$, 使当 $y \geqslant y_0$ 时， $y\mathbf{P}(|X_1| > y) < \delta$, 于是对一切 $n > y_0$

$$
\begin{aligned} \frac{\mathbf{E}X_{n1}^2}{2n} &\leqslant \frac{1}{n}\int_0^{y_0} y\mathbf{P}(|X_1| > y)dy + \frac{1}{n}\int_{y_0}^{n} y\mathbf{P}(|X_1| > y)dy \\ &\leqslant \frac{y_0}{n}\sup_{0\leqslant y\leqslant y_0}[y\mathbf{P}(|X_1| > y)] + \delta. \end{aligned}
$$

由此即知 $\lim_{n\to\infty} \mathbf{E}X_{n1}^2/n = 0$, 因而由 (3)-(4) 知 $S_n/n - a_n \xrightarrow{\mathbf{P}} 0$.

必要性的证明用到对称化的手法．为此，我们先给出

*1).**定义** 若 r.v. X 与 $-X$ 同分布，则称它为 **对称** r.v.；若 r.v. X_1, X_2 独立，且都与 X 同分布，则称 $\dot{X} := X_1 - X_2$ 为 X 的一个 **对称化** r.v. (它是一个对称 r.v.), 而 $\dot{F} := F_{X_1 - X_2} = F_X^{*2}$

称为 F_X 的 **对称化分布**. 显然由乘积测度定理容易由 X 出发构造出它的对称化 r.v.. 下面要进行更一般的构造.

设 $X = \{X_n : n \in \mathbb{N}\}$ 为一独立 r.v. 序列, 由乘积测度定理存在两个相互独立的 r.v. 序列 $X^k = \{X_n^k : n \in \mathbb{N}\}$, $k = 1, 2$, 并且都与 X 同分布. 于是 $\dot{X} := \{\dot{X}_n := X_n^1 - X_n^2 : n \in \mathbb{N}\}$ 仍然是独立 r.v. 序列, 并且它的每一分量是 X 的相应分量的对称化. 这种构造 (X^1, X^2) 的手法, 实际上是一种 (独立) 耦合.

现在证明两条关于对称化手法有用的引理:

*2). **引理** (对称化不等式) 设 X_1, X_2 是独立同分布的 r.v., 则

$$\mathbf{P}(|X_1 - X_2| > u) \leqslant 2\mathbf{P}(|X_1| > \frac{1}{2}u), \quad \forall u > 0. \tag{5}$$

若 $a \geqslant 0$ 满足 $\mathbf{P}(X_1 \leqslant a) \geqslant p$ 及 $\mathbf{P}(X_1 \geqslant -a) \geqslant p$, 则

$$\mathbf{P}(|X_1 - X_2| > u) \geqslant p\mathbf{P}(|X_1| > u + a), \quad \forall u > 0, \tag{6}$$

特别若 0 是 X 的中数, 则有

$$\mathbf{P}(|X_1 - X_2| > u) \geqslant \frac{1}{2}\mathbf{P}(|X_1| > u). \quad \forall u > 0 \tag{7}$$

证明 因为对任何 $u > 0$, 必有

$$\begin{aligned}
&\{\omega : |X_1(\omega)| > \frac{1}{2}u\} \cup \{\omega : |X_2(\omega)| > \frac{1}{2}u\} \\
&\supset \{\omega : |X_1(\omega) - X_2(\omega)| > u\} \\
&\supset \{\omega : X_1(\omega) > u + a, X_2(\omega) \leqslant a\} \\
&\cup \{\omega : X_1(\omega) < -u - a, X_2(\omega) \geqslant -a\}
\end{aligned}$$

所以由 X_1, X_2 独立同分布知 (5),(6) 成立. □

*3). **引理** 设 X_k, $k = 1, \cdots, n$, 是对称 r.v. 且独立, 则 $S_n := \sum_{k=1}^n X_k$ 为对称 r.v., 且有

$$\mathbf{P}(|S_n| > u) \geqslant \frac{1}{2}\mathbf{P}(\max_{1 \leqslant k \leqslant n} |X_k| > u) \quad \forall u > 0. \tag{8}$$

若 X_k 还同分布, 则有

$$\mathbf{P}(|S_n| > u) \geqslant \frac{1}{2}(1 - e^{-n\mathbf{P}(|X_1|>u)}. \tag{9}$$

证明 令 $M := X_T$, 其中

$$T(\omega) := \min\{k : |X_k(\omega)| = \max_{1\leqslant i\leqslant n} |X_i(\omega)|\}.$$

再令 $L := S_n - M$, 则易见 (L, M), $(-L, M)$, $(L, -M)$, $(-L, -M)$ 同分布 (留作习题). 于是对任何 $u > 0$ 有

$$\begin{aligned}\mathbf{P}(S_n > u) &= \mathbf{P}(M + L > u) \geqslant \mathbf{P}(M > u, L \geqslant 0) \\ &= \frac{1}{2}[\mathbf{P}(M > u, L \geqslant 0) + \mathbf{P}(M > u, -L \geqslant 0)] \geqslant \frac{1}{2}\mathbf{P}(M > u).\end{aligned}$$

同理 $\mathbf{P}(S_n < -u) \geqslant \frac{1}{2}\mathbf{P}(M < -u)$. 故 (8) 获证.

若 X_k 还同分布, 则由

$$\mathbf{P}(\max |X_k| \leqslant u) = \prod_{k=1}^{n} \mathbf{P}(|X_k| \leqslant u) = [\mathbf{P}(|X_1| \leqslant u)]^n$$

及初等不等式 $1 - u < e^{-u}$. 对一切 $u \in (0, 1)$ 成立, 即得 (9) 式. □

必要性的证明: 象 1) 那样构造对称化独立 r.v. 序列 $\dot{X}$, 它的部分和 $\dot{S}_n$ 是 $S_n - na_n$ 的对称化, 设 m 是 X_1 的中数. 依次应用引理 2) 的 (5), 引理 3) 的 (9) 及引理 2) 的 (6) 即得

$$\begin{aligned}2\mathbf{P}(|S_n - na_n| > n\varepsilon) &\geqslant \mathbf{P}(|\dot{S}_n| > 2n\varepsilon) \\ &\geqslant \frac{1}{2}[1 - \exp(-n\mathbf{P}(|\dot{X}_1| > 2n\varepsilon))] \\ &\geqslant \frac{1}{2}[1 - \exp(-\frac{n}{2}\mathbf{P}(|X_1| > 2n\varepsilon + |m|))].\end{aligned}$$

由弱大数律成立知上式左端当 $n \to \infty$ 时趋于 0, 于是上式右端趋于 0, 因而条件的必要性成立.

Хинчин 定理的证明如下：若 $\mathbf{E}|X_1| < \infty$，则

$$x\mathbf{P}(|X_1| > x) \leqslant \mathbf{E}\left[|X_1| I_{\{|X_1| > x\}}\right] \to 0, \quad \text{当} x \to \infty \text{时},$$

又

$$a_n = \mathbf{E}\left[|X_1| I_{\{|X_1| \leqslant n\}}\right] \to \mathbf{E}|X_1|, \qquad (n \to \infty)$$

所以

$$\frac{S_n}{n} - \mathbf{E}|X_1| = \frac{S_n}{n} - a_n - (\mathbf{E}|X_1| - a_n) \xrightarrow{\mathbf{P}} 0. \quad \square$$

在定理 4 的证明中实际考虑的是 X_{nk}, $k = 1, 2, \cdots, n$, $S'_n = \sum_{k=1}^n X_{nk}$, $n \in \mathbb{N}$, 的大数律 $S'_n/n - a_n \xrightarrow{\mathbf{P}} 0$. 关于这方面我们有下列更一般的定理.

***5. 定理** 对一切 $n \in \mathbb{N}$, X_{nk}, $k = 1, 2, \cdots, k_n$, 是独立 r.v.，且 X_{nk} 的分布函数为 F_{nk}, 令 $b_n \to \infty$, $b_n > 0$, 设 X_{nk} 满足下列条件

(1)
$$\sum_{k=1}^{k_n} \mathbf{P}(|X_{nk}| > b_n) \to 0 \qquad (n \to \infty)$$

及

(2)
$$b_n^{-2} \sum_{k=1}^{k_n} \int_{|x| \leqslant b_n} x^2 \, dF_{nk}(x) \to 0 \qquad (n \to \infty)$$

则

$$\frac{S_n - a_n}{b_n} \xrightarrow{\mathbf{P}} 0 \qquad (n \to \infty),$$

其中 $S_n = \sum_{k=1}^{k_n} X_{nk}$, $a_n = \sum_{k=1}^{k_n} \int_{|x| \leqslant b_n} x \, dF_{nk}(x)$.

证明 令 $S'_n := \sum_{k=1}^{k_n} X_{nk}^{b_n}$ 则其证明步骤与定理 4 一样，首先由 (1)

$$\mathbf{P}(S_n \neq S'_n) \leqslant \sum_{k=1}^{k_n} \mathbf{P}(|X_{nk}| > b_n) \to 0.$$

其次，对一切 $\varepsilon>0$，(注意 $a_n=\mathbf{E}S'_n=\sum_{k=1}^{k_n}\mathbf{E}X_{nk}^{b_n}$) 有

$$\mathbf{P}\left(\frac{|S'_n-a_n|}{b_n}>\varepsilon\right)\leqslant\frac{1}{b_n^2\varepsilon^2}\sigma^2(S'_n)=\frac{1}{b_n^2\varepsilon^2}\sum_{k=1}^{k_n}\sigma^2(X_{nk}^{b_n})$$
$$\leqslant\frac{1}{b_n^2\varepsilon^2}\sum_{k=1}^{k_n}\mathbf{E}(X_{nk}^{b_n})^2\to 0,$$

故定理获证.

这条定理较之定理 4 有两点推广：i) $b_n\neq n$, 这一点可参考 [Ch1] 5.2 练习 5, 6; ii) 不是序列，而是三角阵，这一点可参考 [Du] 例 5.8 (p. 37)，随机排列. 关于 r.v. 的三角阵满足弱大数律的讨论可参看 [LLL] 第二、三章.

习题

1. 设 $X_1,X_2,\cdots,X_n,\cdots$ *iid*, 且 X_1 服从 Cauchy 分布，即

$$\mathbf{P}(X_1\leqslant x)=\int_{-\infty}^{x}\frac{dt}{\pi(1+t^2)},\qquad x\in\mathbb{R}.$$

试证： $x\mathbf{P}(|X_1|>x)\to 2/\pi\neq 0$. 因而由定理知不存在 a_n 使 $S_n/n-a_n\xrightarrow{\mathbf{P}}0$.

2. 设 $X_1,\cdots,X_n,\cdots$ 为 *iid* 序列，且

$$\mathbf{P}(X_n=(-1)^{k-1}k)=\frac{c}{k^2\log k},\qquad k\geqslant 3,$$

其中 c 满足 $\sum_{k=3}^{\infty}\frac{c}{k^2\log k}=1$. 试证 $\mathbf{E}|X_1|=\infty$，但有一常数 a 使 $\frac{S_n}{n}\xrightarrow{\mathbf{P}}a$.

3. 令 $p_k=1/2^kk(k+1)$, $k\in\mathbb{N}$, $p_0=1-(p_1+p_2+\cdots)$, 设 $X_1,X_2,\cdots,X_n,\cdots$ 为 *iid* 序列，

$$\mathbf{P}(X_n=-1)=p_0;\qquad \mathbf{P}(X_n=2^k-1)=p_k,\qquad k\in\mathbb{N},$$

则 $\mathbf{E}X_n=0$, 令 $S_n=X_1+\cdots+X_n$, 试应用定理 5 证明

$$\frac{S_n}{n/\log_2 n}\xrightarrow{\mathbf{P}}-1,$$

其中 $\log_2 n$ 表示 n 的以 2 为底的对数.

4. 证明：定理 4 的证明中的引理 3) 中的 (L,M), $(L,-M)$, $(-L,M)$, $(-L,-M)$ 同分布. (提示：因为 $X_1,\cdots,X_n$ 都是对称 r.v. 且独立，于是对任何 Borel 可测函数 $f(x_1,\cdots,x_n)$ 来说，$f(\widetilde{X}_1,\cdots,\widetilde{X}_n)$, $\widetilde{X}_k=X_k$ 或 $-X_k$, $k=1,\cdots,n$ 都是同分布的.)

§3. 随机级数的收敛

下面我们讨论强大数律成立的条件，为此先讨论随机变量组成的级数的收敛性. 本节假设所遇随机变量都是有限值的.

1. 设 $\{X_n, n\in\mathbb{N}\}$ 是一个 r.v. 序列. 我们考察

$$\sum_{n=1}^{\infty} X_n \text{ 及 } \frac{1}{b_n}\sum_{k=1}^{n} X_k (b_n\to\infty)$$

收敛区域的特性. 记 $\mathcal{T}_n := \sigma(\{X_k: k>n\})$, 即 $\mathcal{T}_n$ 是序列 $\{X_n: n\in\mathbb{N}\}$ 自 n 以后的尾 $\{X_k: k\geqslant n+1\}$ 所生成的 σ–代数，并称 σ–代数

$$\mathcal{T} := \bigcap_{n\in\mathbb{N}} \mathcal{T}_n$$

为 $\{X_n: n\in\mathbb{N}\}$ 的 **尾 σ– 代数**. 尾 σ– 代数中的元称为 **尾事件**. 由于 $\sum_{k=1}^{\infty} X_k$ 与 $\sum_{k=n+1}^{\infty} X_k$ 的收敛域相同，当 $b_n\to\infty$ 时，$(1/b_n)\sum_{k=1}^{n} X_k$ 与 $(1/b_{n+m})\sum_{k=n+1}^{n+m} X_k$ 的收敛域相同，所以对一切 $n\in\mathbb{N}$, $\sum_{k=1}^{\infty} X_k$, $(1/b_n)\sum_{k=1}^{n} X_k$ 的收敛域属于 $\mathcal{T}_n$, 因而属于 $\mathcal{T}$.

关于独立 r.v. 的尾 σ– 代数有下列定理.

2. 定理 (Колмогоров 0–1 律) 任一独立 r.v. 序列 $\{X_n: n\in\mathbb{N}\}$ 的尾事件的概率为 0 或 1.

证明 只需证：对一切 $A\in\mathcal{T}$, $\mathbf{P}(A)^2=\mathbf{P}(A\cap A)=\mathbf{P}(A)$. 亦即 A 与 A 独立.

事实上，对一切 $n \in \mathbb{N}$、$\mathcal{T}_n \supset \mathcal{T}$ 与 $\sigma(X_1, \cdots X_n)$ 独立，所以 $\mathcal{T}$ 与 $\bigcup_{n=1}^{\infty} \sigma(X_1, \cdots, X_n)$ 独立，因而 $\mathcal{T}$ 与 $\sigma(\bigcup_{n=1}^{\infty} \sigma(X_1, \cdots, X_n))$ 独立．但后者 $\supset \mathcal{T}$. 故 $\mathcal{T}$ 与 $\mathcal{T}$ 独立． □

定理 2 虽然说明了独立 r.v. 序列 $\{X_n : n \in \mathbb{N}\}$ 的级数 $\sum_n X_n$ 或 $(1/b_n)\sum_{k=1}^n X_k$, $(b_n \to \infty)$ 的收敛性只有两个极端情形： a.s. 收敛或 a.s. 不收敛．但是它并不能给出判断 a.s. 收敛的具体准则．这方面有用的是下述定理．

3. 定理 (Borel 0-1 律) 设 $\{A_n : n \in \mathbb{N}\}$ 是独立事件列，则

$$\mathbf{P}(A_n \text{ i.o.}) = \begin{cases} 0, & \text{当} \sum\limits_{n=1}^{\infty} \mathbf{P}(A_n) \begin{cases} < \infty, \\ = \infty. \end{cases} \\ 1, & \end{cases}$$

证明 前一结论是 Borel-Cantelli 引理的推论．今往证后一结论．由于 $A_n, n \in \mathbb{N}$ 独立，

$$\begin{aligned} \mathbf{P}(A_n \text{ i.o.}) &= \mathbf{P}\Big(\bigcap_n \bigcup_{k \geqslant n} A_k\Big) \\ &= 1 - \lim_{n\to\infty} \lim_{m\to\infty} \textstyle\prod_{k=n}^m \mathbf{P}(A_k^c) \end{aligned}$$

而

$$\begin{aligned} \textstyle\prod_{k=n}^m \mathbf{P}(A_k^c) &= \textstyle\prod_{k=n}^m (1 - \mathbf{P}(A_k)) \\ &\leqslant \textstyle\prod_{k=n}^m e^{-\mathbf{P}(A_k)} = e^{-\sum_{k=n}^m \mathbf{P}(A_k)}. \end{aligned}$$

故由 $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$ 知 $\mathbf{P}(A_n \text{ i.o.}) = 1$. □

4. 推论 若 $X_n, n \in \mathbb{N}$ 为独立 r.v. 列，则 $X_n \xrightarrow{\text{a.s.}} 0$ 当且仅当对一切 $c \in (0, \infty)$, $\sum_{k=1}^{\infty} \mathbf{P}(|X_k| \geqslant c) < \infty$.

证明 $X_n \xrightarrow{\text{a.s.}} 0$ 当且仅当对一切 $c > 0$,

$$0 = \mathbf{P}\Big(\bigcap_n \bigcup_{k \geqslant n} \{|X_k| \geqslant c\}\Big) = \mathbf{P}(|X_n| \geqslant c \text{ i.o.})$$

因而由定理 3 知推论获证． □

下面来建立独立 r.v. 的级数和 a.s. 收敛的准则，它是建立在 Borel–Cantelli 引理及下列重要的 Колмогоров 不等式的基础上. 可以说各种大数律的研究都是建立在相应的不等式的基础上的.

5. 定理 (Колмогоров 不等式) 设 $\{X_n:\ n\in\mathbb{N}\}$ 是独立 r.v. 列，$\mathbf{E}X_n$ 及 $\sigma^2(X_n)$ 有限，记 $S_n:=\sum_{k=1}^n X_k$，则对一切 $\varepsilon>0$,

$$\mathbf{P}\Big(\max_{1\leqslant k\leqslant n}|S_k-\mathbf{E}S_k|\geqslant\varepsilon\Big)\leqslant\frac{1}{\varepsilon^2}\sum_{k=1}^n\sigma^2(X_k). \tag{1}$$

若还存在 A 使

$$\forall\, n\ \text{有}\ |X_k-\mathbf{E}X_k|\leqslant A<\infty. \tag{2}$$

则还有

$$\mathbf{P}\Big(\max_{1\leqslant k\leqslant n}|S_k-\mathbf{E}S_k|\geqslant\varepsilon\Big)\geqslant 1-\frac{(\varepsilon+A)^2}{\sum_{k=1}^n\sigma^2(X_k)}. \tag{3}$$

(注：(1) 中的事件包含 Чебышев 不等式中的事件，因而 (1) 比后者强.)

证明 不失一般性，可设 $\mathbf{E}X_n=0$ (否则用 $X_n-\mathbf{E}X_n$ 代替 X_n).

固定 $\varepsilon>0$, 对于集

$$\Lambda:=\Big\{\omega:\ \max_{1\leqslant k\leqslant n}|S_k(\omega)|\geqslant\varepsilon\Big\} \tag{4}$$

中的任一 ω, 令

$$\nu(\omega):=\min\big\{k:1\leqslant k\leqslant n,\ |S_k(\omega)|\geqslant\varepsilon\big\}$$

则 ν 是定义在 Λ 上的 r.v. (如果约定 $\min\varnothing=\infty$, 即 $\nu(\omega):=\infty$, 对一切 $\omega\in\Lambda^c$, 则 ν 为可测函数). 令

$$\begin{aligned}\Lambda_k&:=\{\omega:\ \nu(\omega)=k\}\\&=\{\omega:\ \max_{1\leqslant j\leqslant k-1}|S_j(\omega)|<\varepsilon,\quad |S_k(\omega)|\geqslant\varepsilon\},\\&\quad k=1,2,\cdots,n,\end{aligned} \tag{5}$$

其中对 $k=1$, $\max_{1\leqslant j\leqslant 0}|S_j(\omega)|$ 理解为 0. 此处 ν 是所说极大值 $\geqslant\varepsilon$ 的 "首次时间". 而 Λ_k 是事件 "k 是首次时间". (此处 ν 为一个 "**停时**", 停时技巧是现代概率论特别是随机过程论中广泛应用的方法). 于是

$$\Lambda=\bigcup_{k=1}^{n}\Lambda_k, \tag{6}$$

且 Λ_k, $k=1,2,\cdots,n$. 两两不交.

1) 先证不等式 (1). 由于 $\Lambda_k\in\sigma(X_1,\cdots,X_k)$, $S_kI_{\Lambda_k}$ 为 $\sigma(X_1,\cdots,X_k)$ 可测, 而 $S_n-S_k=\sum_{\ell=k+1}^{n}X_\ell$ 为 $\sigma(X_{k+1},\cdots,X_n)$ 可测, 所以 $S_kI_{\Lambda_k}$ 与 S_n-S_k 独立. 于是由 (6) 知

$$\begin{aligned}\int_\Lambda|S_n|^2\,d\mathbf{P}&=\sum_{k=1}^{n}\int_{\Lambda_k}[S_k+(S_n-S_k)]^2\,d\mathbf{P}\\&=\sum_{k=1}^{n}\int_{\Lambda_k}[S_k^2+2S_k(S_n-S_k)+(S_n-S_k)^2]\,d\mathbf{P}.\end{aligned}\tag{7}$$

而由独立 r.v. 的期望性质知

$$\begin{aligned}&\int_{\Lambda_k}S_k(S_n-S_k)\,d\mathbf{P}=\mathbf{E}\big[(S_kI_{\Lambda_k})(S_n-S_k)\big]\\&=\mathbf{E}(S_kI_{\Lambda_k})\mathbf{E}(S_n-S_k)=\mathbf{E}(S_kI_{\Lambda_k})\cdot\sum_{\ell=k+1}^{n}\mathbf{E}X_\ell=0,\end{aligned}\tag{8}$$

故由 (7), (8), (5), (6), (4) 各式知

$$\begin{aligned}\sigma^2(S_n)\geqslant\int_\Lambda S_n^2\,d\mathbf{P}\geqslant\sum_{k=1}^{n}\int_{\Lambda_k}S_k^2\,d\mathbf{P}\geqslant\varepsilon^2\sum_{k=1}^{n}\mathbf{P}(\Lambda_k)\\=\varepsilon^2\mathbf{P}(\Lambda)=\varepsilon^2\mathbf{P}\big(\max_{1\leqslant k\leqslant n}|S_k|\geqslant\varepsilon\big).\end{aligned}$$

此即 (1) 式.

2) 再在假设 (2) 之下证明 (3) 式. 令

$$A_k:=\big\{\max_{1\leqslant j\leqslant k}|S_j|<\varepsilon\big\}, \tag{9}$$

(则 $A_k\downarrow$ 且 $\Lambda_k=A_{k-1}\setminus A_k$). 注意到

$$S_{k-1}I_{A_{k-1}}+X_kI_{A_{k-1}}=S_kI_{A_{k-1}}=S_kI_{A_k}+S_kI_{\Lambda_k},$$

$$S_{k-1}、I_{A_{k-1}} \text{ 都与} X_k \text{ 独立},$$

$$I_{A_k}I_{\Lambda_k}=I_{A_k\cap\Lambda_k}=0, \tag{10}$$

并将 (10) 式两边的绝对值平方, 然后取期望得

$$\begin{aligned}&\mathbf{E}|S_{k-1}I_{A_{k-1}}|^2+\sigma^2(X_k)\mathbf{P}(A_{k-1})\\&=\mathbf{E}|S_kI_{A_k}|^2+\mathbf{E}|S_kI_{\Lambda_k}|^2,\end{aligned} \tag{11}$$

再由 (2) 得

$$|S_kI_{\Lambda_k}|\leqslant|S_{k-1}I_{\Lambda_k}|+|X_kI_{\Lambda_k}|\leqslant(\varepsilon+A)I_{\Lambda_k}. \tag{12}$$

将 (12) 代入 (11), 并注意 $\mathbf{P}(A_{k-1})\geqslant\mathbf{P}(A_n)$, 即得

$$\begin{aligned}&\mathbf{E}|S_{k-1}I_{A_{k-1}}|^2+\sigma^2(X_k)\mathbf{P}(A_n)\\&\leqslant\mathbf{E}|S_kI_{A_k}|^2+(\varepsilon+A)^2\mathbf{P}(\Lambda_k).\end{aligned}$$

对 $k=1,2,\cdots,n$ 求和 (仍然约定 $S_0\equiv0$), 即得

$$\begin{aligned}\sum_{k=1}^{n}\sigma^2(X_k)\mathbf{P}(A_n)&\leqslant\mathbf{E}|S_nI_{A_n}|^2+(\varepsilon+A)^2\sum_{k=1}^{n}\mathbf{P}(\Lambda_k)\\&\leqslant\varepsilon^2\mathbf{P}(A_n)+(\varepsilon+A)^2\mathbf{P}(\Lambda)\leqslant(\varepsilon+A)^2\end{aligned}$$

再注意到 $A_n^c=\Lambda$ 即得 (3) 式. □

6. 引理 设 $\{X_n:\ n\in\mathbb{N}\}$ 为独立 r.v. 序列,

1) 若 $\sum_{n=1}^{\infty}\sigma^2(X_n)<\infty$, 则 $\sum_{n=1}^{\infty}(X_n-\mathbf{E}X_n)$ a.s. 收敛;

2) 若 $\sum_{n=1}^{\infty}\sigma^2(X_n)=\infty$ 且对一切 $n\in\mathbb{N}$, $|X_n-\mathbf{E}X_n|\leqslant A<\infty$, 则 $\sum_{n=1}^{\infty}(X_n-\mathbf{E}X_n)$ a.s. 发散.

证明 $\sum_{n=1}^{\infty}(X_n-\mathbf{E}X_n)$ a.s. 收敛, a.s. 发散就是序列 $S_n-\mathbf{E}S_n$, $n\in\mathbb{N}$, a.s. 收敛, a.s. 发散的问题, 不失一般性可设 $\mathbf{E}S_n=0$.

1) 由 a.s. 收敛的判别条件知为证 1) 只需证明：对一切 $\varepsilon > 0$,

$$\mathbf{P}\Big(\bigcup_k\{|S_{n+k} - S_n| \geqslant \varepsilon\}\Big) \to 0 \quad (n \to \infty). \tag{1}$$

而由 Колмогоров 不等式 (5.1) 及 $\sum_{n=1}^{\infty}\sigma^2(X_n) < \infty$ 知

$$\mathbf{P}\Big(\bigcup_k\{|S_{n+k} - S_n| \geqslant \varepsilon\}\Big) = \mathbf{P}\Big(\bigcup_k\Big\{\max_{1\leqslant\nu\leqslant k}|S_{n+\nu} - S_n| \geqslant \varepsilon\Big\}\Big)$$

$$= \lim_{k\to\infty}\mathbf{P}\Big(\max_{1\leqslant\nu\leqslant k}|S_{n+\nu} - S_n| \geqslant \varepsilon\Big) \leqslant \lim_{k\to\infty}\frac{1}{\varepsilon^2}\sum_{\nu=1}^{k}\sigma^2(X_{n+\nu}) \tag{2}$$

$$= \frac{1}{\varepsilon^2}\sum_{k=n+1}^{\infty}\sigma(X_k) \to 0 \quad (n \to \infty).$$

故 (1) 成立，因而 $\{S_n\}$ a.s. 收敛.

2) 注意到

$$\Big\{\sum X_n \text{ 发散}\Big\} = \bigcup_{\varepsilon>0}\bigcap_{n=1}^{\infty}\bigcup_{\nu=1}^{\infty}\{|S_{n+\nu} - S_n| \geqslant \varepsilon\}$$

$$\supset \bigcap_n\bigcup_\nu\Big\{\max_{1\leqslant k\leqslant\nu}|S_{n+k} - S_n| \geqslant \varepsilon\Big\}$$

而由 Колмогоров 不等式 (5.3) 及 $\sum_{n=1}^{\infty}\sigma^2(X_n) = \infty$, $|X_n| \leqslant A$ 知

$$\mathbf{P}\Big(\bigcup_\nu\Big\{\max_{1\leqslant k\leqslant\nu}|S_{n+k} - S_n| \geqslant \varepsilon\Big\}$$

$$= \lim_{\nu\to\infty}\mathbf{P}\Big(\max_{1\leqslant k\leqslant\nu}|S_{n+k} - S_n| \geqslant \varepsilon\Big)$$

$$\geqslant \lim_{\nu\to\infty}\left(1 - \frac{(\varepsilon + A)^2}{\sum_{k=1}^{\nu}\sigma^2(X_{n+k})}\right) = 1.$$

注意到可数个概率为 1 的事件的交的概率仍为 1, 所以

$$\mathbf{P}(\textstyle\sum X_n\text{发散}) \geqslant \mathbf{P}\Big(\bigcap_n\bigcup_\nu\Big\{\max_{1\leqslant k\leqslant\nu}|S_{n+\nu} - S_n| \geqslant \varepsilon\Big\}\Big) = 1.$$

故 $\sum_{n=1}^{\infty}X_n$a.s. 发散. □

7. 定理 (三级数定理) 设 $\{X_n: n\in\mathbb{N}\}$ 为独立 r.v. 序列，对于一个给定的常数 A, 定义 Y_n 为 X_n 在 A 处的截割，即

$$Y_n(\omega):=\begin{cases}X_n(\omega), & \text{当}|X_n(\omega)|\leqslant A,\\ 0, & \text{当}|X_n(\omega)|>A,\end{cases}$$

则级数 $\sum_n X_n$ a.s. 收敛的充要条件是下面的三级数同时收敛：

(1) $\sum_n \mathbf{P}(|X_n|>A)=\sum_n \mathbf{P}(X_n\neq Y_n)$;

(2) $\sum_n \mathbf{E}Y_n$;

(3) $\sum_n \sigma^2(Y_n)$.

证明 充分性：由于 Y_n, $n\in\mathbb{N}$, 独立且 (3) 收敛，所以由引理 6(1) 知 $\sum_n(Y_n-\mathbf{E}Y_n)$a.s. 收敛，再由 (2) 收敛即知 $\sum_n Y_n$a.s. 收敛. 由 (1) 收敛及推论 2.3 知 $\sum_n X_n$a.s. 收敛.

必要性：由 $\sum_n X_n$a.s. 收敛知 $X_n\xrightarrow{\text{a.s.}}0$, 于是由推论 4 知 $\sum_{k=1}^\infty \mathbf{P}(|X_n|\geqslant A)<\infty$, 即 (1) 收敛. 再由推论 2.3 知 $\sum_n Y_n$a.s. 收敛，而且对一切 $n\in\mathbb{N}$, $|Y_n-\mathbf{E}Y_n|\leqslant 2A<\infty$, 在 $\Omega\times\Omega$ 上定义：

$$Y_n^{(1)}(\omega_1,\omega_2):=Y_n(\omega_1),$$
$$Y_n^{(2)}(\omega_1,\omega_2):=Y_n(\omega_2),$$

则 $Y_n^{(1)}$, $Y_n^{(2)}$, $n\in\mathbb{N}$, 独立. 令

$$Z_n:=Y_n^{(2)}-Y_n^{(1)},$$

则 $\mathbf{E}Z_n=0$, $|Z_n|\leqslant 2A$, $\sum Z_n$ a.s. 收敛， Z_n, $n\in\mathbb{N}$, 独立，故由引理 6(2) 知

$$2\sum\sigma^2(Y_N)=\sum\sigma^2(Z_n)<\infty$$

即 (3) 收敛. 再由引理 6.1) 知 $\sum_n(Y_n-\mathbf{E}Y_n)$a.s. 收敛. 前面已证 $\sum_n Y_n$a.s. 收敛，所以 (2) 收敛. □

由定理 8.2.4 知，在有限测度下，可测函数序列 a.s. 收敛蕴含依测度收敛，反之不然，但对独立 r.v. 的级数来说，a.s. 收敛与依概率收敛是一致的，这就是下列的定理.

8. 定理 若 $\{X_n:\ n\in\mathbb{N}\}$ 是独立 r.v. 序列，则 $\sum_n X_n$ a.s. 收敛与依概率收敛等价.

***证明**：只需证 $\sum_n X_n$ 依概率收敛蕴含 $\sum X_n$ a.s. 收敛.

由 $\sum_n X_n$ 依概率收敛知对一切 $\varepsilon>0$, 存在 n_ε, 使

$$\mathbf{P}(|S_{n+k}-S_n|>\varepsilon)<\varepsilon,\quad \forall k\geqslant 1, n\geqslant n_\varepsilon, \tag{1}$$

而 $\sum_n X_n$ a.s. 收敛的充分必要条件是：存在 $\varepsilon_m\downarrow 0(m\to\infty)$, 使

$$\lim_{m\to\infty}\lim_{n\to\infty}\lim_{k\to\infty}\mathbf{P}\Big(\max_{1\leqslant\nu\leqslant k}|S_{n+\nu}-S_n|\geqslant 2\varepsilon_m\Big)=0. \tag{2}$$

易见有

$$\begin{aligned}\bigcup_{\nu=1}^{k}\Big\{\max_{1\leqslant j\leqslant\nu-1}|S_{n+j}-S_n|\leqslant 2\varepsilon,\ |S_{n+\nu}-S_n|>2\varepsilon,&\\ |S_{n+k}-S_{n+\nu}|\leqslant\varepsilon\Big\}&\\ \subset\Big\{|S_{n+k}-S_n|>\varepsilon\Big\}&\end{aligned}$$

而且上式左端各项两两不交．再由 $\Big\{\max\limits_{1\leqslant j\leqslant\nu-1}|S_{n+j}-S_n|\leqslant 2\varepsilon,$ $|S_{n+\nu}-S_n|>2\varepsilon\Big\}$ 与 $\{|S_{n+k}-S_{n+\nu}|\leqslant\varepsilon\}$ 的独立性知

$$\begin{aligned}&\mathbf{P}\Big(\max_{1\leqslant\nu\leqslant k}|S_{n+\nu}-S_n|>2\varepsilon\Big)\min_{1\leqslant\nu\leqslant k}\mathbf{P}(|S_{n+k}-S_{n+\nu}|\leqslant\varepsilon)\\ &\leqslant\sum_{\nu=1}^{k}\mathbf{P}\Big(\max_{1\leqslant j\leqslant\nu-1}|S_{n+j}-S_n|\leqslant 2\varepsilon,\ |S_{n+\nu}-S_n|>2\varepsilon\Big)\\ &\qquad\mathbf{P}(|S_{n+k}-S_{n+\nu}|\leqslant\varepsilon)\\ &=\mathbf{P}\Big(\bigcup_{\nu=1}^{k}\Big\{\max_{1\leqslant j\leqslant\nu-1}|S_{n+j}-S_n|\leqslant 2\varepsilon,\ |S_{n+\nu}-S_n|>2\varepsilon,\end{aligned}$$

$$|S_{n+k} - S_{n+\nu}| \leqslant \varepsilon\}\Big)$$
$$\leqslant \mathbf{P}(|S_{n+k} - S_n| > \varepsilon)$$

于是由 (1) 知

$$\mathbf{P}\Big(\max_{1\leqslant\nu\leqslant k} |S_{n+\nu} - S_n| > 2\varepsilon\Big) < \frac{\varepsilon}{1-\varepsilon}, \quad \forall\, n \geqslant n_\varepsilon, \quad k \geqslant 1.$$

用 (6.2) 相同算法即得：对一切 $n \geqslant n_\varepsilon$

$$\mathbf{P}\Big(\bigcup_k\{|S_{n+k} - S_n| > 2\varepsilon\}\Big)$$
$$= \lim_{k\to\infty} \mathbf{P}\Big(\max_{1\leqslant\nu\leqslant k} |S_{n+\nu} - S_n| > 2\varepsilon\Big) < \frac{\varepsilon}{1-\varepsilon}$$

故对一切 $m \in \mathbb{N}$,

$$\limsup_{n\to\infty} \mathbf{P}\Big(\bigcup_k\{|S_{n+k} - S_n| > 2\varepsilon_m\}\Big) < \frac{\varepsilon_m}{1-\varepsilon_m}.$$

再令 $m \to \infty$ 即得 (2), 即 S_n, $n \in \mathbb{N}$, a.s. 收敛. □

习题

1. 设 X_n, $n \in \mathbb{N}$, 为 *iid* 序列, 且 $\mathbf{P}(X_n = -1) = \mathbf{P}(X_n = 1) = 1/2$, 则级数 $\sum_n (X_n/(n^\theta))$ 当 $\theta \in (1/2, 1]$ 时 a.s. 收敛, 当 $\theta \in [0, 1/2]$ 时 a.s. 发散.

2. 证明推广的 Колмогоров 不等式：若 $X_n, n \in \mathbb{N}$, 独立, $r \geqslant 1, \mathbf{E}X_n = 0$, 对 $c > 0$, 记

$$C = \{\max_{1\leqslant k\leqslant n} |S_k| \geqslant c\},$$

则

$$c^r \mathbf{P}(C) \leqslant \mathbf{E}(|S_n|^r I_C) \leqslant \mathbf{E}(|S_n|^r),$$

应用此不等式证明：若 $S_n \xrightarrow{r} S$(有限), 则 $S_n \xrightarrow{\text{a.s.}} S$.

§4. 强大数律

首先证明将级数的收敛性与强大数律联系起来的下述引理.

1. 引理 (Kronecher) 设 $\{x_k : k \in \mathbb{N}\} \subset \mathbb{R}$, $0 < a_k \uparrow \infty$, $(k \in \mathbb{N})$, 则

$$\sum_n \frac{x_n}{a_n} \text{ 收敛} \Longrightarrow \frac{1}{a_n}\sum_{k=1}^n x_k \to 0. \tag{1}$$

证明 令 $a_0 := 0$, $b_0 := 0$ 且

$$b_n := \sum_{k=1}^n \frac{x_k}{a_k}, \qquad n \in \mathbb{N}.$$

则 $x_n = a_n(b_n - b_{n-1})$, $n \in \mathbb{N}$, 于是

$$\begin{aligned}\frac{1}{a_n}\sum_{k=1}^n x_k &= \frac{1}{a_n}\left\{\sum_{k=1}^n a_k b_k - \sum_{k=1}^n a_k b_{k-1}\right\}\\ &= \frac{1}{a_n}\left\{a_n b_n - \sum_{k=1}^n (a_k - a_{k-1})b_{k-1}\right\}\\ &= b_n - \sum_{k=1}^n \frac{a_k - a_{k-1}}{a_n} b_{k-1}\end{aligned} \tag{2}$$

由 $a_0 = 0$, $a_k - a_{k-1} \geqslant 0$, 知

$$\sum_{k=1}^n \frac{a_k - a_{k-1}}{a_n} = 1. \tag{3}$$

又由假设 $b_\infty := \lim_{n\to\infty} b_n$ 有限, 由此可证:

$$\sum_{k=1}^n \frac{a_k - a_{k-1}}{a_n} b_{k-1} \to b_\infty, \tag{4}$$

于是由 (2), (4) 即得 $(1/a_n)\sum_{k=1}^n x_k \to 0$ $(n \to \infty)$.

至于 (4) 的证明如下: 由假设知 $B := \sup_n |b_n| \in (0, \infty)$ (若 $B = 0$, 则一切平凡). 对一切 $\varepsilon > 0$, 取 $K \in \mathbb{N}$, 使当 $k \geqslant K$ 时, $|b_k - b_\infty| < \varepsilon/2$, 取 $N \in \mathbb{N}$, 使当 $n \geqslant N$ 时, $a_K/a_n < \varepsilon/(4B)$. 于

是当 $n > \max\{N, K\}$ 时，由 (3) 及上述取法即得

$$\begin{aligned}
&\left|\sum_{k=1}^{n}\frac{(a_k - a_{k-1})}{a_n}b_{k-1} - b_\infty\right| \\
&\leqslant \left(\sum_{k=1}^{K} + \sum_{k=K+1}^{n}\right)\frac{a_k - a_{k-1}}{a_n}|b_{k-1} - b_\infty| \\
&\leqslant 2B\sum_{k=1}^{K}\frac{a_k - a_{k-1}}{a_n} + \frac{\varepsilon}{2}\sum_{k=K+1}^{n}\frac{a_k - a_{k-1}}{a_n} \\
&\leqslant 2B\frac{a_K}{a_n} + \frac{\varepsilon}{2} < \varepsilon.
\end{aligned}$$

故 (4) 获证. □

由引理 1 及引理 3.6 即得下述定理.

2. 定理 设 $\{Z_n: n \in \mathbb{N}\}$ 是独立 r.v. 列，$\mathbf{E}Z_n = 0$, $T_n = Z_1 + \cdots + Z_n$, 若 $0 < a_n \uparrow \infty$, 且 $\sum_{n=1}^{\infty}(\mathbf{E}Z_n^2/a_n^2) < \infty$, 则 $T_n/a_n \to 0$, a.s.

证明 由引理 3.6.1) 知 $\sum_n(Z_n/a_n)$ a.s. 收敛，所以由引理 1 知 $(T_n/a_n) \xrightarrow{\text{a.s.}} 0$.

3. 注 1) 可以证明如下的比引理 3.6 更广的结果：设 $\varphi: \mathbb{R} \to \mathbb{R}_+ \setminus \{0\}$, $\varphi(x) = \varphi(-x)$, 且当 $|x| \uparrow$ 时，$(\varphi(x)/|x|) \uparrow$, $(\varphi(x)/x^2) \downarrow$. 若 $\{X_n: n \in \mathbb{N}\}$ 为独立 r.v. 序列，$\mathbf{E}X_n = 0$, $0 < a_n \uparrow \infty$, 且

$$\sum_n \frac{\mathbf{E}\varphi(X_n)}{\varphi(a_n)} < \infty,$$

则 $\sum_n(X_n/a_n)$ a.s. 收敛.

在此结果中取 $\varphi(x) = x^2$, 即为引理 3.6. 还有一些其它结论可参看 [Ch1] §5.4.

2) 设 $X_n, n \in \mathbb{N}$, 为 *iid* r.v. 列，且 $\mathbf{E}X_n = \mu$, 对 $Z_n := X_n - \mu$, $a_n := n$ 应用定理 2 即得 $\mathbf{E}X_n^2 < \infty$ 情形的强大数律. 但是我们可以适当选取 a_n 使结果大大加强.

令 $a_n = n^{1/2}[\log(n \vee e)]^{1/2+\varepsilon}$, $\varepsilon > 0$ 任意给定，则由于

$$\sum_{n=N}^{\infty}\left(\frac{1}{\sqrt{n}[\log n]^{1/2+\varepsilon}}\right)^2 = \sum_{n=N}^{\infty}\frac{1}{n(\log n)^{1+2\varepsilon}}$$

$$= O\left(\int_N^{\infty}\frac{dx}{x(\log x)^{1+2\varepsilon}}\right) = O\left(\int_N^{\infty}\frac{d\log x}{(\log x)^{1+2\varepsilon}}\right)$$

$$= O\left(\frac{1}{2\varepsilon(\log N)^{2\varepsilon}}\right) = 0(1), \qquad N \to \infty,$$

有 $\sum_n \dfrac{\sigma^2(X_n)}{n\log(n \vee e)^{1+2\varepsilon}} < \infty$. 因而由定理 2 知

$$\frac{S_n - n\mu}{n^{1/2}[\log(n \vee e)]^{1/2+\varepsilon}} \xrightarrow{\text{a.s.}} 0.$$

另一方面，可以证明

$$\varlimsup_{n\to\infty}\frac{S_n - n\mu}{[n\log\log(n \vee e^e)]^{1/2}} = c > 0. \quad \text{a.s.}$$

所以上述结果距最好结果不远.

下面是独立同分布 r.v. 的强大数律，它是 Хинчин 弱大数律在强大数律中的对照物.

4. 定理 (Колмогоров) 设 $\{X_n : n \in \mathbb{N}\}$ 是 *iid* r.v. 列，则

1) $\mathbf{E}|X_1| < \infty \implies S_n/n \to \mathbf{E}X_1$, a.s.

2) $\mathbf{E}|X_1| = +\infty \implies \varlimsup_{n\to\infty}|S_n/n| = +\infty$, a.s.

由 1)、2) 可推出下列结论：$(S_n/n) \xrightarrow{\text{a.s.}}$(某一有限常数)$c$ 的充要条件是 $\mathbf{E}|X_1| < \infty$; 在条件成立时，$c = \mathbf{E}X_1$.

证明 先证 1) 令 $Y_n = X_n I_{\{|X_n| \leqslant n\}}$, 则

$$\sum_n \mathbf{P}(X_n \neq Y_n) = \sum_n \mathbf{P}(|X_n| > n) = \sum \mathbf{P}(|X_1| > n)$$

$$\leqslant \int_0^{\infty}\mathbf{P}(|X_1| > x)\,dx = \mathbf{E}|X_1| < \infty.$$

则由推论 2.3 知若能证明 $(1/n)\sum_{k=1}^n Y_k \xrightarrow{\text{a.s.}} \mathbf{E}X_1$, 则 1) 获证.

现在应用定理 2 来证明： $(1/n)\sum_{k=1}^n(Y_k-\mathbf{E}Y_k)\xrightarrow{\text{a.s.}}0$, 为此只需证 $\sum_{n=1}^\infty\dfrac{\sigma^2(Y_n)}{n^2}<\infty$. 而

$$\sum_{n=1}^\infty\frac{\sigma^2(Y_n)}{n^2}\leqslant\sum_{n=1}^\infty\frac{\mathbf{E}Y_n^2}{n^2}=\sum_n\frac{1}{n^2}\int_{|x|\leqslant n}x^2\,dF(x)$$

$$=\sum_{n=1}^\infty\frac{1}{n^2}\sum_{j=1}^n\int_{j-1<|x|\leqslant j}x^2dF(x)=\sum_{j=1}^\infty\int_{j-1<|x|\leqslant j}x^2\,dF(x)\sum_{n=j}^\infty\frac{1}{n^2}$$

$$\leqslant\sum_{j=1}^\infty j\int_{j-1<|x|\leqslant j}|x|dF(x)\frac{c}{j}=c\int_{-\infty}^\infty|x|dF(x)=c\mathbf{E}|X_1|<\infty$$

其次由控制收敛定理知

$$\mathbf{E}Y_k=\mathbf{E}X_kI_{\{|X_k|\leqslant k\}}=\mathbf{E}X_1I_{\{|X_1|\leqslant k\}}\xrightarrow{k\to\infty}\mathbf{E}X_1,$$

所以

$$(1/n)\sum_{k=1}^n\mathbf{E}Y_k\xrightarrow{k\to\infty}\mathbf{E}X_1.$$

故得

$$\frac{1}{n}\sum_{k=1}^nY_k\xrightarrow{\text{a.s.}}\mathbf{E}X_1.$$

因而 1) 获证.

今往证 2): 对一切 $A>0$,

$$+\infty=\mathbf{E}\frac{|X_1|}{A}=\int_0^\infty\mathbf{P}\Big(\frac{|X|}{A}>x\Big)\,dx$$

$$=\sum_{n=1}^\infty\int_{n-1}^n\mathbf{P}\Big(\frac{|X_1|}{A}>x\Big)dx\leqslant\sum_{n=1}^\infty\mathbf{P}(|X_1|>A(n-1))$$

故

$$\sum_{n=1}^\infty\mathbf{P}(|X_n|\geqslant An)=+\infty.$$

由于 $\{|X_n|\geqslant An\}$, $n\in\mathbb{N}$, 为独立事件，所以由 Borel 0–1 律知

$$\mathbf{P}(\{|X_n|\geqslant An\}\text{ i.o.})=1.$$

但 $|S_n - S_{n-1}| = |X_n| \geqslant An$ 蕴含 $|S_n| \geqslant An/2$ 或 $|S_{n-1}| \geqslant An/2$, 所以有

$$\mathbf{P}\Big(\frac{|S_n|}{n} \geqslant \frac{A}{2} \text{ i.o.}\Big) = 1,$$

即

$$\overline{\lim_{n\to\infty}} \frac{|S_n|}{n} \geqslant \frac{A}{2} \quad \text{a.s.}$$

由于 $A > 0$ 任意, 所以 2) 获证.

当 $X_n, n \in \mathbb{N}$, *iid* 且 $\mathbf{E}|X_1| = \infty$ 时, Feller 有一关于强大数律的推广, 见 [Ch1] 5.4.3(p.138). 限于篇幅不再介绍.

习题

1. 若 $\{X_n\}$ 为 *iid* 序列, $\mathbf{E}X_1^+ = \infty, \mathbf{E}X_1^- < \infty$, 则 $S_n/n \to \infty$ a.s.. 进一步, 只要 $\mathbf{E}X_1$ 存在, 就有 $S_n/n \to \mathbf{E}X_1$. a.s..

2. 应用定理 3 证明: 若 $\{X_n\}$ 为独立随机变量序列,

1) 设存在 $p \in [1, 2]$, 使 $\sum_n \frac{1}{n^p}\mathbf{E}(|X_n|^p) < \infty$, 则 $\frac{1}{n}\sum_1^n X_j \to 0$ a.s.

2) 若对某个 $\delta \in (0, 1]$ 和 $M < \infty$, 使对一切 n 有 $\mathbf{E}|X_n|^{1+\delta} \leqslant M$, 则 $\frac{1}{n}\sum_1^n X_j \to 0$ a.s.

3. 若 $\{X_n\}$ 为 *iid* 序列, 且 $\mathbf{E}X_1 = 0$, $\{c_n\}$ 是一个有界实数序列, 则

$$\frac{1}{n}\sum_1^n c_j X_j \to 0 \text{ a.s.}$$

§5. 应 用

现在我们介绍几个应用.

1. 关于概率和数学期望 (及矩) 的计算.

设有一 r.v. X, $\mathbf{E}|X|<\infty$, 对母体独立地抽取容量为 n 的样本，设为 $X_1,\cdots,X_n$, 则 $X_1,\cdots,X_n$ 可以认为是独立且与 X 同分布的 r.v. 当 $n\to\infty$ 时，由强大数律.

$$\frac{\sum_{k=1}^n X_k}{n}\xrightarrow{\text{a.s.}}\mathbf{E}X. \tag{1}$$

这就告诉我们：如果第 k 次抽到的值是 x_k, 则按强大数律就有：当 n 充分大时，$(\sum_{k=1}^n x_k)/n$, 几乎可以看作是 $\mathbf{E}X$ 的近似值.

注意这并不能算作 $(\sum_{k=1}^n x_k)/n$ 近似 $\mathbf{E}X$ 的证明，而只是一种解释. 让我们看最简单的情形. $X:=I_A$, 于是对 A 作 n 次独立试验，就得到独立同分布 r.v. I_{A_k}, $k=1,2,\cdots,n$, 而 $\sum_{k=1}^n I_{A_k}=\nu_n(A)$ 为 A 在 n 次独立试验中出现的频数，于是由强大数律 (设 $\mathbf{P}(A)=p$)

$$\frac{\nu_n(A)}{n}\xrightarrow{\text{a.s.}}p=\mathbf{E}I_A. \tag{2}$$

这似乎是证明了“频率 → 概率”. 其实不然. 为说明这一点，我们取 $\Omega=\{0,1\}$, $\mathbf{P}(\{1\})=p=1-\mathbf{P}(\{0\})$, $A=\{1\}$, 而 (2) 的含义是指

$$\widetilde{\mathbf{P}}\left(\frac{\nu_n(A)}{n}\to p\right)=1$$

其中 $\widetilde{\mathbf{P}}=\prod_{k=1}^\infty \mathbf{P}_n$, $\mathbf{P}_n(\{0\})=1-p$, $\mathbf{P}_n(\{1\})=p$. 因而 $\widetilde{\mathbf{P}}(\cdot)$ 的统计解释有赖于 $\mathbf{P}(\cdot)$ 的统计解释，即有赖于 p 的统计解释，所以并未证明命题“频率 → 概率”.

但是不管怎样，(1) 式还是可以作为估计期望的理论解释. 通常统计书就是这样做的，而且当 $\mathbf{E}|X|^\ell<\infty$ 时，

$$\frac{\sum_{k=1}^n X_k^\ell}{n}\xrightarrow{\text{a.s.}}\mathbf{E}X^\ell,$$

就是矩法的理论根据.

2. 和 1.(2) 有密切关系的问题是 "**经验分布**" 的问题.

设 X 是一 r.v., 且 $X \sim F$, F 未知, 用统计术语说 X 是一未知母体. 对母体 X 独立地抽取样品, 得 $X_1, X_2, \cdots$, 则可设 X_k, $k \in \mathbb{N}$, 为 *iid* r.v. 序列, 且 $X_k \sim F$, 而对一切 $\omega \in \Omega$, $X_k(\omega)$, $k \in \mathbb{N}$, 称为母体 X 的观察值.

给定 $n \in \mathbb{N}$ 和 $\omega \in \Omega$, 将 $X_1(\omega), \cdots, X_n(\omega)$ 按递增顺序排列为

$$Y_{n1}(\omega) \leqslant Y_{n2}(\omega) \leqslant \cdots \leqslant Y_{nn}(\omega), \tag{1}$$

(有很多统计书上记 Y_{nk} 为 X_k^*!), 再定义

$$F_n(x, \omega) := \begin{cases} 0, & x < Y_{n1}(\omega); \\ \dfrac{k}{n}, & Y_{nk}(\omega) \leqslant x < Y_{n,k+1}(\omega); \quad x \in \mathbb{R} \\ 1, & Y_{nn}(\omega) \leqslant x. \end{cases} \tag{2}$$

即对一切 $x \in \mathbb{R}$, $F_n(x, \omega)$ 为使 $X_j(\omega) \leqslant x$, $j = 1, 2, \cdots, n$, 发生的频率, 即 n 次独立试验中事件 $\{X \leqslant x\}$ 的频率, 而函数 $F_n(\cdot, \omega)$ 为取自母体 X(或 F) 的 n 个样本值的 **经验分布函数**. (实际上它是一个随机过程 — 通常称随机过程 $\sqrt{n}(F_n(x, \omega) - F(x))$, $x \in \mathbb{R}$, $\omega \in \Omega$, 为经验过程).

令

$$\xi_k(x) := I_{\{X_k \leqslant x\}}, \qquad k = 1, 2, \cdots,$$

则对一切 $x \in \mathbb{R}$, $\xi_k(x)$, $k \in \mathbb{N}$, 为 *iid* 序列,

$$\xi_k(x) \sim \begin{pmatrix} 0 & 1 \\ \mathbf{P}(X > x) & \mathbf{P}(X \leqslant x) \end{pmatrix}$$

且

$$F_n(x) = \frac{1}{n} \sum_{k=1}^{n} \xi_k(x),$$

$\mathbf{E}\xi_k(x) = F(x)$, 于是由强大数律知: 对一切 $x \in \mathbb{R}$

$$F_n(x, \cdot) \xrightarrow{\text{a.s.}} F(x). \tag{3}$$

如果我们只关心对特定的 x 的 $F(x)$, 那么讨论就完成了. 但是由于 x 可以是 $\mathbb{R}$ 中的任何值, 所以我们自然希望能够得到同时关于一切 $x \in \mathbb{R}$ 的结果.

首先我们注意 (3) 的精确意义是: 对一切 $x \in \mathbb{R}$, 存在 $N(x) \in \mathcal{F}$, 使 $\mathbf{P}(N(x)) = 0$, 且

$$(4) \qquad \forall\, \omega \notin N(x), \quad F_n(x,\omega) \to F(x) \quad (n \to \infty),$$

于是对 $\mathbb{R}$ 的任何可数子集 Q(例如有理数集), 令 $N := \bigcup_{x\in Q} N(x)$, 则 $\mathbf{P}(N) = 0$, 且

$$(5) \quad \forall\, \omega \in \Omega \setminus N = N^c, \quad F_n(x,\omega) \to F(x) \quad (n \to \infty), \quad x \in Q.$$

实际上可以证明上述收敛可以对一切 x 而且是一致的, 即下述定理.

3. 定理 (Гливенко)

$$(1) \qquad \mathbf{P}\Big(\sup_{-\infty<x<\infty} |F_n(x) - F(x)| \to 0\Big) = 1$$

证明 设 $J := \{x \in \mathbb{R}:\ F(x) - F(x-) > 0\}$, 则 J 为可数集, 对一切 $x \in J$, 令

$$\eta_k(x,\omega) := \begin{Bmatrix} 1, & X_k(\omega) = x \\ 0, & X_k(\omega) \neq x \end{Bmatrix} = I_{\{X_k=x\}}(\omega).$$

于是对一切 $x \in J$,

$$F_n(x,\omega) - F_n(x-,\omega) = \frac{1}{n}\sum_{k=1}^{n} \eta_k(x,\omega).$$

因此和 2 的讨论一样, 由强大数律知: 存在 $N(x)$ 使 $\mathbf{P}(N(x)) = 0$ 且

$$(2) \quad F_n(x,\omega) - F_n(x-,\omega) \to F(x) - F(x-), \qquad \forall\, \omega \notin N(x).$$

令 Q 为有理数集, 并令

$$N_1 := \bigcup_{x\in Q\cup J} N(x).$$

则 $\mathbf{P}(N_1)=0$ 且对一切 $\omega \in N_1^c$, 有

$$
(3)\qquad \begin{cases} F_n(x,\omega)-F_n(x-,\omega)\to F(x)-F(x-), & \forall x\in J,\\ F_n(x,\omega)\to F(x). & \forall\, x\in Q, \end{cases}
$$

于是定理由 (3) 及下列的纯分析引理即得.

引理 设 F_n 与 F 为右连续概率分布函数, Q, J 如前所述且

$$
(4)\qquad \begin{aligned} &F_n(x)\to F(x), \qquad \forall\, x\in Q,\\ &F_n(x)-F_n(x-)\to F(x)-F(x-), \qquad \forall\, x\in J. \end{aligned}
$$

则

$$
\sup_{-\infty<x<\infty}|F_n(x)-F(x)|\to 0.
$$

证明 设引理不成立, 即 $\overline{\lim}_{n\to\infty}\sup_{-\infty<x<\infty}|F_n(x)-F(x)|=2\varepsilon>0$, 则存在 $\{n_k\}\subset\mathbb{N}$ 及 $\{x_k\}\subset\mathbb{R}$ 使

$$
|F_{n_k}(x_k)-F(x_k)|\geqslant\varepsilon
$$

由 F_n, F 为概率分布函数知 $x_k\not\to\pm\infty$, 因此有 $\{x_k\}$ 的子序列 (不妨设就是它本身) 使

$$
x_k\to\xi\in\mathbb{R}.
$$

由于 $\{x_k\}$ 是无穷序列, 所以或者有无穷多个 $x_k<\xi$, 或者有无穷多个 $x_k\geqslant\xi$. 因此不妨设 (1) $x_k\uparrow\xi$, $x_k<\xi$ 或 (2) $x_k\downarrow\xi$, $x_k\geqslant\xi$, 又 $|F_{n_k}(x_k)-F(x_k)|$ 也是无穷多项, 和上面讨论一样, 不妨设 $|F_{n_k}(x_k)-F(x_k)|=F_{n_k}(x_k)-F(x_k)$, 对一切 $k\in\mathbb{N}$, 或 $|F_{n_k}(x_k)-F(x_k)|=F(x_k)-F_{n_k}(x_k)$, 对一切 $k\in\mathbb{N}$, 所以可以分以下四种情形讨论: 先任取 $r_1, r_2\in Q$, $r_1<\xi<r_2$.

1) $x_k\uparrow\xi$, $x_k<\xi$, $|F_{n_k}(x_k)-F(x_k)|=F_{n_k}(x_k)-F(x_k)$, 此时

$$
\varepsilon\leqslant F_{n_k}(x_k)-F(x_k)\leqslant F_{n_k}(\xi-)-F(r_1), \quad (\text{当 } k \text{ 充分大})
$$

$$
\begin{cases}
\leqslant F_{n_k}(r_2) - F(r_1) \xrightarrow{k\to\infty} F(r_2) - F(r_1) \\
\xrightarrow[r_1\uparrow\xi]{r_2\downarrow\xi} F(\xi) - F(\xi) = 0, & \text{当}\xi \notin J, \\
\leqslant F_{n_k}(\xi-) - F_{n_k}(\xi) + F_{n_k}(r_2) - F(r_1) \\
\xrightarrow{k\to\infty} F_k(\xi-) - F(\xi) + F(r_2) - F(r_1) \xrightarrow[r_1\uparrow\xi]{r_2\downarrow\xi} 0, & \text{当}\xi \in J.
\end{cases}
$$

2)$x_k \uparrow \xi$, $x_k < \xi$, $|F_{n_k}(x_k) - F(x_k)| = F(x_k) - F_{n_k}(x_k)$, 此时

$$
\begin{aligned}
\varepsilon \leqslant F(x_k) - F_{n_k}(x_k) &\leqslant F(\xi-) - F_{n_k}(r_1)l \\
&\to F(\xi-) - F(r_1) \xrightarrow[r_1\uparrow\xi]{} 0.
\end{aligned}
$$

(第二个不等式要求 k 充分大使得 $x_k > r_1$ 成立)

3) $x_k \downarrow \xi$, $x_k \geqslant \xi$,

$$
\varepsilon \leqslant F(x_k) - F_{n_k}(x_k) \leqslant F(r_2) - F_{n_k}(\xi)
$$

$$
\begin{cases}
\leqslant F(r_2) - F_{n_k}(r_1) \xrightarrow{k\to\infty} F(r_2) - F(r_1) \\
\xrightarrow[r_1\uparrow\xi]{r_2\downarrow\xi} F(\xi) - F(\xi), & \text{当}\xi \notin J, \\
\leqslant F(r_2) - F_{n_k}(r_1) + F_{n_k}(\xi-) - F_{n_k}(\xi) \\
\xrightarrow{k\to\infty} F(r_2) - F(r_1) + F(\xi-) - F(\xi) \xrightarrow[r_1\uparrow\xi]{r_2\downarrow\xi} 0, & \text{当}\xi \in J.
\end{cases}
$$

4) $x_k \downarrow \xi$, $x_k \geqslant \xi$,

$$
F_{n_k}(x_k) - F(x_k) \leqslant F_{n_k}(r_2) - F(\xi) \xrightarrow{k\to\infty} F(r_2) - F(\xi) \xrightarrow{r_2\downarrow\xi} 0.
$$

无论哪种情形都得出矛盾结果： $\varepsilon \leqslant 0$. 故引理成立.

*4. 应用于 **更新理论**. 设 $\{X_n : n \in \mathbb{N}\}$ 是 *iid* r.v. 列. 且 $X_n \geqslant 0$, 而不几乎处处等于零, 即 $\mathbf{P}(X_n = 0) < 1$, 我们可以将 X_n 看成是在一个设备中所使用 (不断更新) 的一个个元件的 "寿命", 也可以看成一代一代生物的年令, 或依次服务序列的每一服务的时间等等, 总之可以解释成经受一个更新过程的某种对象的

"寿命" 或某些循环现象重新出现的周期. 这里有许多需要探讨的实际问题和理论问题. 例如: 1) 直到时刻 t, 有多少次更新? 2) 最后一次更新在多久以前? 3) 下一次更新在多久以后, 等等.

我们来考虑第一问题: 令

$$(1)\qquad N(t,\omega) := \sup\{n:\ S_n(\omega) \leqslant t\}, \qquad \forall\, t \geqslant 0,\quad \omega \in \Omega,$$

即 $N(t)$ 表示在 $[0,t]$ 内更新次数, 显然 (约定 $S_0 \equiv 0$).

$$(2)\quad \{\omega:\ N(t,\omega) = n\} = \{\omega:\ S_n(\omega) \leqslant t < S_{n+1}(\omega)\}, \quad \forall\, n \geqslant 0.$$

对 $n \leqslant m-1$ 求和即得

$$\begin{aligned}(3)\qquad \{\omega:\ N(t,\omega) < m\} &= \bigcup_{n=0}^{m-1}\{\omega:\ N(t,\omega) = n\}\\ &= \bigcup_{n=0}^{m-1}\{\omega:\ S_n(\omega) \leqslant t < S_{n+1}(\omega)\}\\ &= \{\omega:\ S_m(\omega) > t\}\end{aligned}$$

故对一切 $t \geqslant 0$, $N(t)$ 是一个 r.v. , 其值取自然数, 而 r.v. 族 $\{N(t):\ t \in [0,\infty)\}$ 称为一个 **更新过程** (renewal process). 若 $X_n \sim E(\lambda)$, 即对一切 $t \geqslant 0$, $\mathbf{P}(X_n \leqslant t) = 1 - e^{-\lambda t}$ $(\lambda > 0)$, 则上述过程即称为简单 Poisson 过程, (或称 Poisson 过程). 首先有

5. 命题 更新的总次数随时间无限增长而趋于无穷大, 即

$$(1)\qquad \lim_{t\to\infty} N(t) = +\infty, \quad \text{a.s.}(\mathbf{P}).$$

证明 由于对一切 $\omega \in \Omega$, $N(t,\omega)\uparrow$, (当 $t\uparrow$) 因而 $\lim_{t\to\infty} N(t) = \sup_{0\leqslant t<\infty} N(t)$, 对一切 $\omega \in \Omega$ 存在 (显然它可测). 若 (1) 不成立, 则

$$\mathbf{P}\left(\sup_{0\leqslant t<\infty} N(t) < \infty\right) > 0,$$

从而存在 $M \in \mathbb{N}$, 使

$$\mathbf{P}\left(\sup_{0\leqslant t<\infty} N(t) < M\right) > 0,$$

于是由 (4.3) 知对一切 $t>0$,

$$\mathbf{P}(S_M>t)=\mathbf{P}(N(t)<M)\geqslant\mathbf{P}\left(\sup_{t\in\mathbb{R}_+}N(t)<M\right)>0.$$

故

$$\mathbf{P}(S_M=+\infty)>0.$$

由于每一 r.v. X_n 有限，所以上式是不可能的. □

其次有

6. 命题 设 $0<m:=\mathbf{E}X_1<\infty$, 则

$$\lim_{t\to\infty}\frac{S_{N(t)}}{N(t)}=m,\ \text{a.s.}\tag{1}$$

证明 由强大数律知存在 N_1, 使 $\mathbf{P}(N_1)=0$,

$$\forall\,\omega\in\Omega\setminus N_1,\qquad\frac{S_n(\omega)}{n}\to m\quad(n\to\infty)$$

由命题 5, 存在 N_2 使 $\mathbf{P}(N_2)=0$, 对一切 $\omega\in\Omega\setminus N_2$, 有

$$\lim_{t\to\infty}N(t,\omega)=\infty,$$

于是对一切 $\omega\in(N_1\cup N_2)^c$, 有

$$\lim_{t\to\infty}\frac{S_{N(t,\omega)}(\omega)}{N(t,\omega)}=m.$$

而 $\mathbf{P}(N_1\cup N_2)=0$, 所以 (1) 成立. □

由 $N(t)$ 的定义知 $S_{N(t)}\leqslant t<S_{N(t)+1}$, 因此 $S_{N(t)}$ 与 t 是接近的，于是自然想到有下列更强的结果.

7. 定理 设 $m:=\mathbf{E}X_1$, 且约定 $1/\infty=0$, 则有

$$\lim_{t\to\infty}\frac{N(t)}{t}=\frac{1}{m},\ \text{a.s.}\tag{1}$$

及

$$\lim_{t\to\infty}\frac{\mathbf{E}N(t)}{t}=\frac{1}{m}.\tag{2}$$

证明 由

$$S_{N(t,\omega)} \leqslant t < S_{N(t)+1}, \ \forall\, \omega \in \Omega,$$

知当 t 充分大 $(t > X_1(\omega))$ 时, $N(t,\omega) > 0$, 因而有

$$\frac{S_{N(t,\omega)}(\omega)}{N(t,\omega)} \leqslant \frac{t}{N(t,\omega)} < \frac{S_{N(t,\omega)+1}(\omega)}{N(t,\omega)+1} \cdot \frac{N(t,\omega)+1}{N(t,\omega)}.$$

对一切 $\omega \in (N_1 \cup N_2)^c$. ($N_1, N_2$ 如第 6 目定义!), 则当 $m < \infty$ 时 (1) 式成立, 当 $m = \infty$ 时, 只需注意此时 $\mathbf{E}X_1 = \infty$, $X_1 \geqslant 0$, 即可证 $S_n/n \to +\infty$ a.s. 因而 (1) 对 $m = +\infty$ 也成立.

为证 (2), 我们注意 $\mathbf{P}(X_n > 0) > 0$, 故存在 $\delta > 0$ 使

$$\forall\, n \in \mathbb{N}, \qquad \mathbf{P}(X_n \geqslant \delta) =: p > 0$$

令

$$X'_n := \delta I_{\{X_n \geqslant \delta\}}$$

对应于 $\{X'_n:\ n \in \mathbb{N}\}$, 作 S'_n, $N'(t)$. 显然有 $S'_n \leqslant S_n$, $N'(t) \geqslant N(t)$, 对一切 $t \geqslant 0$. 又因 $\{X'_n/\delta\}$ 独立且

$$X'_n/\delta \sim \begin{pmatrix} 0 & 1 \\ \mathbf{P}(X_n<\delta) & \mathbf{P}(X_n \geqslant \delta) \end{pmatrix},$$

于是经计算得

$$\mathbf{E}\{N'(t)^2\} = O\left(\frac{t^2}{\delta^2}\right), \qquad t \to \infty.$$

因此对固定的 δ, 我们有

$$\mathbf{E}\left\{\left(\frac{N(t)}{t}\right)^2\right\} \leqslant \mathbf{E}\left\{\left(\frac{N'(t)}{t}\right)^2\right\} = O(1).$$

因为 (1) 蕴含 $N(t)/t$ 依概率收敛于 $1/m$, 故取 $X_n = N(n)/n$, $r = 2$, 应用推论 8.3.19.2) 就得出 (2) 中 t 被 n 代替的情形, 由此即可得出 (2) 本身. □

第十章 特征函数和中心极限定理

特征函数是在概率论中有着重要应用的分析工具，其原因在于：

1° 它与分布函数一一对应，由特征函数的某些性质可以求得分布律的相应性质，而且在分布律的弱收敛与特征函数的收敛之间也保持着对应关系，因而可通过求特征函数的极限求得随机变量序列的极限分布；

2° 特征函数是有界一致连续函数，比分布函数更易于应用分析的工具；

3° 独立随机变量和的特征函数是各被加项特征函数的乘积，计算和研究都很方便，从而在历史上推动了古典极限问题的解决.

§1. 特 征 函 数 的 定 义 及 简 单 性 质

1. 定义 设 X 是 $(\Omega, \mathcal{F}, \mathbf{P})$ 上的实随机变量，其分布函数为 F，概率分布为 μ，则称 $\mathbb{R}$ 上的复值函数

$$
\begin{aligned}
f_X(u) := \mathbf{E}e^{iuX} &= \int_{-\infty}^{\infty} e^{iux}\, dF(x) \\
&= \int_{\mathbb{R}} e^{iux}\, d\mu = \int_{\mathbb{R}} \cos(ux)\, d\mu + i\int_{\mathbb{R}} \sin(ux)\, d\mu
\end{aligned}
$$

为 r.v. X(相应地分布函数 F 或概率分布 μ) 的 **特征函数**.

由定义 1 容易得到下述

2. 命题 特征函数 f 具有下述简单性质：

1) $|f(u)| \leqslant f(0) = 1$, 对一切 $u \in \mathbb{R}$;

2) f 在 $\mathbb{R}$ 上一致连续，即

$$|f(u+h)-f(u)| \leqslant \int_{\mathbb{R}} |e^{ixh}-1|\,\mu(dx) \to 0, \text{当 } h \to 0 \text{ 时;}$$

3) $\overline{f_X(u)} = f_X(-u) = f_{-X}(u)$， 对一切 $u \in \mathbb{R}$;

4) $f_{aX+b} = f_X(au)e^{ibu}$， 对一切 $u \in \mathbb{R}$， $a, b \in \mathbb{R}$;

5) 若 X, Y 独立，则 $f_{X+Y}(u) = f_X(u)f_Y(u)$， 对一切 $u \in \mathbb{R}$.

3. 常见分布律特征函数的例：

例1 设 $X \sim b(n,p)$, 即 $\mathbf{P}(X=k) = \binom{n}{k}p^k(1-p)^{n-k}$, $k = 0, 1, \cdots, n$, 则对一切 $u \in \mathbb{R}$,

$$\begin{aligned} f_X(u) &= \sum_{k=1}^{n} \binom{n}{k} p^k (1-p)^{n-k} e^{iuk} \\ &= \binom{n}{k}(pe^{iu})^k(1-p)^{n-k} = (1-p+pe^{iu})^n. \end{aligned}$$

例2 设 $X \sim \mathcal{P}(\lambda)$, 即 $\mathbf{P}(X=n) = e^{-\lambda}\frac{\lambda^n}{n!}, n \in \mathbb{Z}_+$, 则对一切 $u \in \mathbb{R}$,

$$\begin{aligned} f_X(u) &= \sum_{n=0}^{\infty} e^{-\lambda}\frac{\lambda^n}{n!}e^{iun} \\ &= \sum_{n=0}^{\infty} e^{-\lambda}\frac{(\lambda e^{iu})^n}{n!} = e^{\lambda(e^{iu}-1)}. \end{aligned}$$

例3 设 X_r 服从负二项分布，即第 r 次成功在第 $r+k$ 次出现的概率

$$\mathbf{P}(X_r = r+k) = \binom{r+k-1}{r-1}p^{r-1}q^k p = \binom{-r}{k}p^r(-q)^k, \ k \in \mathbb{Z},$$

其中 p 是每次试验成功的概率， $q = 1-p$, 而

$$\binom{-r}{k} = \frac{-r(-r-1)\cdots(-r-k+1)}{k!} = (-1)^k\binom{r+k-1}{r-1},$$

则对一切 $u \in \mathbb{R}$,

$$f_{X_r}(u) = \sum_{k=0}^{\infty} \binom{-r}{k} p^r (-q)^k e^{iu(r+k)}$$

$$= (pe^{iu})^r \sum_{k=0}^{\infty} \binom{-r}{k} (-qe^{iu})^k = (pe^{iu})^r \Big(\frac{1}{1-qe^{iu}}\Big)^r.$$

后一等式可用 Taylor 级数验证.

例4　设 X 在 [0.1] 上均匀分布，则

$$f_X(u) = \int_a^b e^{iux} \frac{1}{b-a}\, dx = \begin{cases} \frac{e^{ibu}-e^{iau}}{i(b-a)u}, & u \neq 0; \\ 1, & u = 0. \end{cases}$$

例5　设 $X \sim N(0,1)$, 则

$$f_X(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} e^{-\frac{x^2}{2}}\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-iu)^2}{2}} e^{\frac{(iu)^2}{2}}\, dx$$

$$= e^{-\frac{u^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-iu)^2}{2}}\, dx.$$

上式最后的积分为

$$\lim_{x\to\infty} \int_{-x-iu}^{x-iu} e^{-\frac{z^2}{2}}\, dz.$$

由围道积分得

$$\int_{-x-iu}^{x-iu} e^{-\frac{z^2}{2}}\, dz = \int_{-x}^{x} e^{-\frac{z^2}{2}}\, dz$$

$$+ \int_{x}^{x-iu} e^{-\frac{z^2}{2}}\, dz + \int_{-x-iu}^{-x} e^{-\frac{z^2}{2}}\, dz.$$

当 $x \to \infty$ 时，上式右端第一项的极限为 $\sqrt{2\pi}$. 对给定的 x, 在复平面集合 $\{x+iy:\ 0 \leqslant |y| \leqslant |u|\}$ 上，$\Big|e^{-\frac{z^2}{2}}\Big| \leqslant e^{-\frac{x^2}{2}}$, 因而有

$$\lim_{x\to\infty} \left| \int_{x}^{x-iu} e^{-\frac{z^2}{2}}\, dz \right| \leqslant \lim_{x\to\infty} |u| e^{-\frac{x^2}{2}} \to 0.$$

同理,

$$\lim_{x\to\infty}\left|\int_{-x-iu}^{-x} e^{-\frac{z^2}{2}}\,dz\right| \leqslant \lim_{x\to\infty}|u|e^{-\frac{x^2}{2}} \to 0.$$

故得

$$f_X(u)=e^{-\frac{u^2}{2}},\quad \forall u\in\mathbb{R}.$$

若 $X\sim N(a,\sigma^2)$, 由命题 2.4) 易知

$$f_X(u)=e^{iau-\frac{(\sigma u)^2}{2}},\quad \forall u\in\mathbb{R}.$$

例6 若 X 是自由度为 n 的 χ^2 变量，即 X 的分布密度为

$$p(x)=\begin{cases}\frac{1}{2^{n/2}\Gamma(n/2)}e^{-x/2}x^{n/2-1}, & \text{当 } x\geqslant 0\text{时};\\ =0, & \text{当 } x<0\text{时}.\end{cases}$$

由围道积分可得

$$f_X(u)=(1-2iu)^{-\frac{n}{2}},\quad \forall u\in\mathbb{R}.$$

4. 命题 设 $f(u)$ 是实 r.v. X 的特征函数，$f^{(k)}(u)$ 表示 $f(u)$ 的 k 阶导数.

1) 若 $f^{(2n)}(0)$ 存在且有限，则对任一满足 $0\leqslant r\leqslant 2n$ 的实数 r, 都有 $\beta_r=\mathbf{E}|X|^r<\infty$.

2) 若 X 的 n 阶绝对矩 $\beta_n=\mathbf{E}|X|^n<\infty$, 则对一切 $k\leqslant n$, 恒有

$$f^{(k)}(u)=i^k\int_{-\infty}^{\infty}x^k e^{iux}\,dF_X(x),\quad u\in\mathbb{R}. \tag{1}$$

因而

$$\alpha_k=\mathbf{E}X^k=i^{-k}f^{(k)}(0). \tag{2}$$

证明 先证明一个分析求导公式：若 $f^{(n)}(0)$ 存在，则

$$f^{(n)}(0)=\lim_{h\to 0}\frac{1}{(2h)^n}\sum_{k=0}^{n}\binom{n}{k}(-1)^k f\big((n-2k)h\big). \tag{3}$$

事实上利用 Taylor 公式我们有

$$f(x) = \sum_{j=0}^{n} f^{(j)}(0)\frac{x^j}{j!} + o(|x|^n).$$

于是

$$\begin{aligned}
&\frac{1}{(2h)^n}\sum_{k=0}^{n}\binom{n}{k}(-1)^k f\big((n-2k)h\big)\\
&= \frac{1}{(2h)^n}\sum_{k=0}^{n}\binom{n}{k}(-1)^k\left[\sum_{j=0}^{n} f^{(j)}(0)\frac{(n-2k)^j h^j}{j!} + o(|(n-2k)h|^n)\right]\\
&= \frac{1}{(2h)^n}\sum_{j=0}^{n} f^{(j)}(0)\frac{h^j}{j!}\sum_{k=0}^{n}\binom{n}{k}(-1)^k(n-2k)^j + o(1).
\end{aligned}$$

利用恒等式

$$\sum_{k=0}^{n}\binom{n}{k}(-1)^k e^{(n-2k)t} = (e^t - e^{-t})^n,$$

两边求 j 阶导数 $(j \leqslant n)$, 再令 $t = 0$ 可证

$$\sum_{k=0}^{n}\binom{n}{k}(-1)^k(n-2k)^j = \begin{cases} 0, & 0 \leqslant j < n. \\ 2^n \cdot n!, & j = n. \end{cases}$$

即得 (3).

再应用 (3) 式求 X 的特征函数在零点的导数. 若 $f^{(m)}(0)$ 存在, 则

$$\begin{aligned}
f^{(m)}(0) &= \lim_{h\to 0}\frac{1}{(2h)^m}\sum_{k=0}^{m}\binom{m}{k}(-1)^k f\big((m-2k)h\big)\\
&= \lim_{h\to 0}\frac{1}{(2h)^m}\int_{-\infty}^{\infty}\sum_{k=0}^{m}\binom{m}{k}(-1)^k e^{i(m-2k)hx}\,dF(x)\\
&= \lim_{h\to 0}\frac{1}{(2h)^m}\int_{-\infty}^{\infty}(e^{ihx} - e^{-ihx})^m\,dF(x)\\
&= \lim_{h\to 0} i^m\int_{-\infty}^{\infty}\left(\frac{\sin(hx)}{hx}\right)^m x^m\,dF(x).
\end{aligned}$$

若 $f^{(2n)}(0)$ 存在, 有限. 则

$$\infty > |f^{(2n)}(0)| = \lim_{h\to 0}\int_{-\infty}^{\infty}\left(\frac{\sin(hx)}{hx}\right)^{2n} x^{2n}\,dF(x)$$
$$\geqslant \int_{-\infty}^{\infty}\varliminf_{h\to 0}\left(\frac{\sin(hx)}{hx}\right)^{2n} x^{2n}\,dF(x)$$
$$= \int x^{2n}\,dF(x) = \beta_{2n}.$$

式中不等号成立是由于 Fatou 引理. 于是 1) 获证.

反之, 若 $\mathbf{E}|X|^n < \infty$, $r \leqslant n$, 则由引理 8.3.8 知

$$\left(\mathbf{E}|X|^r\right)^{1/r} \leqslant \left(\mathbf{E}|X|^n\right)^{1/n} < \infty.$$

所以, 对一切 $k \leqslant n$, 有

$$\int_{-\infty}^{\infty}|x|^k\,dF(x) = \mathbf{E}|X|^k < \infty.$$

而

$$f(u) = \int_{-\infty}^{\infty} e^{iux}\,dF(x).$$

今往证: 当 $\mathbf{E}|X|^k < \infty$ 时, $f^{(k)}(u)$ 存在且

$$f^{(k)}(u) = i^k\int_{-\infty}^{\infty} x^k e^{iux}\,dF(x).$$

事实上, 当 $k = 1$ 时, 由推论 5.4.10 及

$$\left|\frac{de^{iux}}{du}\right| = |ixe^{iux}| = |x|$$

关于 $\mathbf{P}_X$ 可积, $\int_{-\infty}^{\infty} e^{iux}\,dF(x)$ 有限, 知

$$f'(u) = \int_{-\infty}^{\infty} ixe^{iux}\,dF(x).$$

若 (1) 式对 $k-1$ 成立, 即 $f^{(k-1)}(u) = i^{k-1}\int_{-\infty}^{\infty} x^{k-1}e^{iux}\,dF(x)$ 存在, 有限, 则由

$$\left|\frac{d(i^{k-1}x^{k-1}e^{iux})}{du}\right| = |x^k e^{iux}| = |x^k|$$

关于 $\mathbf{P}_X$ 可积知

$$f^k(u) = i^k \int_{-\infty}^{\infty} x^k e^{iux}\, dF(x).$$

从而有

$$\mathbf{E}X^k = i^{-k} f^{(k)}(0). \qquad \square$$

习题

1. 试求均值为 $\dfrac{1}{\lambda}$ 的指数分布的特征函数.

2. 试求在 $[-a, a]$ 上分布的三角分布，即分布密度 $p(x) = \dfrac{a-|x|}{a^2}$, 的特征函数.

3. 如果 f_k, $k = 1, \cdots, n$ 是特征函数， $\lambda_k > 0$, $k = 1, \cdots, n$, $\sum_{k=1}^{n} \lambda_k = 1$, 证明 $\sum_{k=1}^{n} \lambda_k f_k$ 也是特征函数.

4. 试由特征函数的定义，用视察法找出随机变量的分布律，使它的特征函数是下述相应的函数：

i) e^{iau};

ii) $\cos u$;

iii) $\cos^2 u$;

iv) $\sum_{k=0}^{\infty} \lambda_k e^{ia_k u}$, $\lambda_k \geqslant 0$, $\sum_{k=0}^{\infty} \lambda_k = 1$.

v) $\sum_{k=0}^{\infty} \lambda_k \cos ku$, $\lambda_k \geqslant 0$, $\sum_{k=0}^{\infty} \lambda_k = 1$.

vi) $\dfrac{1}{1+iu}$; (提示：利用习题 1 及命题 2.)

5. 试证：若 f 是特征函数，则 $|f|^2$ 也是特征函数. (提示：构造独立随机变量 X_1, X_2, 使得 X_1 与 X 同分布， X_2 与 $-X$ 同分布.)

6. 设 X 的 n 阶绝对矩有限，试证

$$\mathbf{E}(X - \mathbf{E}X)^n = i^{-n} \frac{d^n}{du^n} \left[e^{-iu(\mathbf{E}X)} f_X(u) \right]_{u=0}.$$

7. 试证：如果 $\{f_n\}$ 是 $\mathbb{R}$ 上特征函数序列，对一切 $u\in\mathbb{R}$，$f_n(u)\to g(u)$，且 g 在零点处连续，则 g 在 $\mathbb{R}$ 上一致连续.

§2. 逆转公式及连续性定理

为了证明特征函数唯一决定分布函数，需要计算下述 Dirichlet 积分.

1. 引理 记符号函数为

$$\operatorname{sgn}\alpha=\begin{cases}1, & 若\ \alpha>0,\\ 0, & 若\ \alpha=0,\\ -1, & 若\ \alpha<0.\end{cases}$$

则有

$$(1)\qquad 0\leqslant \operatorname{sgn}\alpha\int_0^y\frac{\sin\alpha x}{x}\,dx\leqslant\int_0^\pi\frac{\sin x}{x}\,dx,\quad \forall\ y\geqslant 0;$$

$$(2)\qquad \lim_{T\to\infty}\int_0^T\frac{\sin\alpha x}{x}\,dx=\frac{\pi}{2}\operatorname{sgn}\alpha;\quad (指\ \mathrm{R}\ 积分)$$

$$(3)\qquad \lim_{T\to\infty}\int_0^T\frac{1-\cos\alpha x}{x^2}\,dx=\frac{\pi}{2}|\alpha|.\quad (指\ \mathrm{R}\ 积分)$$

证明 由 $\operatorname{sgn}\alpha$ 的定义及变量替换公式知只需对 $\alpha=1$ 证明引理.

1) 对一切 $y>0$，存在 n，使 $n\pi\leqslant y<(n+1)\pi$. 于是

$$\int_0^y\frac{\sin x}{x}\,dx=\sum_{k=0}^{n-1}\int_{k\pi}^{(k-1)\pi}\frac{\sin x}{x}\,dx+\int_{n\pi}^y\frac{\sin x}{x}\,dx.$$

由于

$$\begin{aligned}\int_{k\pi}^{(k-1)\pi}\frac{|\sin x|}{x}\,dx&=\int_0^\pi\frac{|\sin x|}{x+k\pi}\,dx\\ &\leqslant\int_{(k-1)\pi}^{k\pi}\frac{|\sin x|}{x}\,dx=\int_0^\pi\frac{|\sin x|}{x+(k-1)\pi}\,dx,\end{aligned}$$

所以 $\sum_{k=0}^{\infty}\int_{k\pi}^{(k+1)\pi}\frac{\sin x}{x}\,dx$ 为各项绝对值递减的交错级数，从而 (1) 式成立且 (2) 中极限存在.

2) 由于

$$\int_0^T \frac{\sin x}{x}\,dx = \int_0^T \sin x \int_0^\infty e^{-ux}\,du\,dx \tag{4}$$

及

$$\int_0^T \int_0^\infty |\sin x| e^{-ux}\,du\,dx = \int_0^T \frac{|\sin x|}{x}\,dx \leqslant T < \infty,$$

因而可用 Fubini 定理交换 (4) 中的积分次序. 由分部积分法易得

$$\int_0^T e^{-ux}\sin x\,dx = \frac{1}{1+u^2}\left[1 - e^{-uT}(u\sin T + \cos T)\right].$$

于是,

$$\begin{aligned}\int_0^T \frac{\sin x}{x}\,dx &= \int_0^\infty \frac{1}{1+u^2}\,du - \int_0^\infty \frac{e^{-uT}}{1+u^2}(u\sin T + \cos T)\,du\\ &= \frac{\pi}{2} - \int_0^\infty \frac{e^{-s}}{T^2+s^2}(s\sin T + T\cos T)\,ds.\end{aligned} \tag{5}$$

当 $T \geqslant 1$ 时,

$$\left|\frac{e^{-s}}{T^2+s^2}(s\sin T + T\cos T)\right| \leqslant 2e^{-s}.$$

上述不等式右端在 $[0,\infty)$ 上 Lebesgue 可积, 而左端当 $T\to\infty$ 时极限为零. 于是 (5) 式右端积分当 $T\to\infty$ 时极限为零, 即 (2) 式获证.

3) 由分部积分法

$$\int_0^T \frac{1-\cos x}{x^2}\,dx = -\frac{1-\cos x}{x}\Big|_0^T + \int_0^T \frac{\sin x}{x}\,dx,$$

再由

$$\lim_{x\to 0}\frac{1-\cos x}{x}=0;\quad \lim_{T\to\infty}\frac{1-\cos T}{T}=0$$

知 (3) 可归结为 (2). □

2. 定理 设 $f(u)$ 是 $(\mathbb{R},\mathcal{B})$ 上概率测度 μ 的特征函数, 则对一切 $x_1<x_2$, 有

$$\mu\big((x_1,x_2)\big)+\frac{1}{2}\mu(\{x_1\})+\frac{1}{2}\mu(\{x_2\})$$

$$=\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-iux_1}-e^{-iux_2}}{iu}f(u)\,du. \tag{1}$$

从而 $(\mathbb{R},\mathcal{B})$ 上概率测度由其特征函数唯一决定 (称 (1) 为 **逆转公式**).

证明 由于

$$\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-iux_1}-e^{-iux_2}}{iu}\int_{-\infty}^{\infty}e^{iux}\,\mu(dx)\,du$$

$$=\int_{-\infty}^{\infty}\int_{-T}^{T}\frac{e^{iu(x-x_1)}-e^{iu(x-x_2)}}{2\pi iu}\,du\,\mu(dx)$$

$$=\int_{-\infty}^{\infty}\left[\frac{1}{\pi}\int_0^T\frac{\sin\big(u(x-x_1)\big)}{u}-\frac{1}{\pi}\int_0^T\frac{\sin\big(u(x-x_2)\big)}{u}\,du\right]\mu(dx)$$

上式方括号中的函数有界, 因而可以在积分号下取极限. 由引理 1 可知, 上式当 $T\to\infty$ 时极限为

$$\int_{-\infty}^{\infty}\left[\frac{\operatorname{sgn}(x-x_1)}{2}-\frac{\operatorname{sgn}(x-x_2)}{2}\right]\mu(dx)$$

$$=\int_{-\infty}^{\infty}\left[I_{(x_1,x_2)}+\frac{1}{2}I_{\{x_1\}}+\frac{1}{2}I_{\{x_2\}}\right]\mu(dx)$$

$$=\mu\big((x_1,x_2)\big)+\frac{1}{2}\mu\big(\{x_1\}\big)+\frac{1}{2}\mu\big(\{x_2\}\big).$$

如果 f 还是概率分布 ν 的特征函数，则当

$$x_1, x_2 \in \Lambda = \{y:\ \mu(\{y\}) = 0, \nu(\{y\}) = 0\}$$

时，

$$\nu((x_1, x_2)) = \mu((x_1, x_2)).$$

令 x_1 沿 Λ 趋于 $-\infty$, 即得 ν 和 μ 的分布函数在它们的共同连续点上相同. 由于不连续点至多可数，所以 ν 和 μ 相同，即特征函数唯一决定概率分布. □

3. 定理 若概率分布函数 F 的特征函数 $f \in L^1(\mathbb{R})$, 即

$$\int_{-\infty}^{\infty} |f(u)|\, du < \infty,$$

则 F 连续可微，且

$$F'(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ixu} f(u)\, du \tag{1}$$

是 F 的密度函数.

证明 在定理 2 的 (1) 式中，令 $x_2 = x, x_1 = x - h, (h > 0)$, 由 $f \in L^1(\mathbb{R})$ 及定理 2 得

$$\begin{aligned}&\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{iuh} - 1}{iu} e^{-iux} f(u)\, du \\ &= \frac{F(x) + F(x-)}{2} - \frac{F(x-h) + F(x-h-)}{2}.\end{aligned}$$

上式中的被积函数被 $|hf|$ 所控制，由控制收敛定理可以在积分号下令 $h \to 0$ 得出右边的极限为 0 的结论. 于是 F 左连续，因而在 $\mathbb{R}$ 中连续. 现在我们可以写出

$$\frac{F(x) - F(x-h)}{h} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{iuh} - 1}{iuh} e^{-iux} f(u)\, du.$$

和前面一样的论证表明，当 $h \to 0$ 时极限存在. 所以 F 在 x 处具有等于 (1) 式右端的左导数，后者显然是连续的. 同理可证 F 具有用同一公式给出的右导数.

由于 $F'(x)$ 连续，所以我们有

$$F(x) = \int_{-\infty}^{x} F'(t)\,dt, \quad \forall x \in \mathbb{R}. \qquad \square$$

4. 定理 **(连续性定理)** 若 F_n, F 为概率分布函数，其相应的特征函数为 f_n, f, 则 $F_n \Longrightarrow F$ 的充分必要条件是对一切 $u \in \mathbb{R}, f_n(u) \to f(u)$.

证明 条件的必要性：记与 F_n、F 相应的概率测度为 μ_n、μ, 由 $F_n \Longrightarrow F$ 有 $\mu_n \Longrightarrow \mu$. 因为对每一 $u \in \mathbb{R}$, e^{iux} 为 $\mathbb{R}$ 上有界连续函数，故由弱收敛的定义知

$$f_n(u) = \int_{-\infty}^{\infty} e^{iux}\,\mu_n(dx) \to \int_{-\infty}^{\infty} e^{iux}\,\mu(dx) = f(u) \quad \forall u \in \mathbb{R}.$$

条件的充分性：由 Fubini 定理，当 $u > 0$ 时，

$$\begin{aligned}
\frac{1}{u}\int_{-u}^{u} (1 - f_n(t))\,dt &= \int_{-\infty}^{\infty} \frac{1}{u}\int_{-u}^{u} \left(1 - e^{itx}\right) dt\mu_n(dx) \\
&= 2\int_{-\infty}^{\infty} \big(1 - \frac{\sin(ux)}{ux}\big)\,\mu_n(dx) \\
&\geqslant 2\int_{|x| \geqslant \frac{2}{u}} \big(1 - \frac{1}{|ux|}\big)\,\mu_n(dx) \\
&\geqslant \mu_n\big(\{x : |x| \geqslant \frac{2}{u}\}\big)
\end{aligned}$$

由于 $f(t)$ 在 $t = 0$ 处连续，因而任给 $\varepsilon > 0$, 存在 $u > o$, 使得

$$\frac{1}{u}\int_{-u}^{u} (1 - f(t))\,dt < \varepsilon.$$

再由对一切 $t \in \mathbb{R}$, $f_n(t) \to f(t)$ 及控制收敛定理知：存在 n_0, 当

$n \geqslant n_0$ 时,

$$\mu_n(\{x : |x| > \frac{2}{u}\}) \leqslant \frac{1}{u}\int_{-u}^{u}(1 - f_n(t))\,dt < \varepsilon.$$

由引理 8.5.11 知 $(\mathbb{R}, \mathscr{B})$ 上每一概率测度都是胎紧的, 对于 $n < n_0$, 存在紧集 K_n, 使得 $\mu_n(K_n^c) < \varepsilon$, 令

$$K = (\bigcup_{n=1}^{n_0-1} K_n) \cup [-u, u],$$

则 K 是紧集且有

$$\mu_n(K^c) < \varepsilon, \qquad \forall\, n \in \mathbb{N}.$$

从而 $\{\mu_n, n \in \mathbb{N}\}$ 是胎紧的. 再由定理 8.5.12 知 $\{\mu_n, n \in \mathbb{N}\}$ 相对紧, 即 $\{\mu_n, n \in \mathbb{N}\}$ 的任一子列都有弱收敛的子列. 设 $\{\mu_{n_k}, n \in \mathbb{N}\}$ 是 $\{\mu_n, n \in \mathbb{N}\}$ 的任一弱收敛的子列, $\mu_{n_k} \Longrightarrow \nu$, 由 $f_{n_k} \to f$ 知 ν 的特征函数也是 f. 从特征函数唯一决定测度, 可得 $\nu = \mu$. 即 $\{\mu_n, n \in \mathbb{N}\}$ 的任一弱收敛的子列有同一弱收敛的极限. 最后由定理 8.5.13 知 $\mu_n \Longrightarrow \mu$. □

5. 推论 若概率测度序列 $\{\mu_n\}$ 相应的特征函数序列 $f_n(u) \to g(u)$, 对一切 $u \in \mathbb{R}$ 成立, 且 g 在 $u = 0$ 处连续, 则存在概率测度 μ, 使得 $\mu_n \Longrightarrow \mu$ 且 μ 的特征函数为 g.

证明 由定理 4 的证明知, g 在 $u = 0$ 处连续蕴含 $\{\mu_n\}$ 的胎紧性, 从而由 Prohorov 定理知它们相对紧. 设 $\{\mu_{n_k}, n \in \mathbb{N}\}$ 是 $\{\mu_n, n \in \mathbb{N}\}$ 的任一弱收敛的子列, $\mu_{n_k} \Longrightarrow \mu$. 再由定理 4 知: μ_{n_k} 的特征函数 f_{n_k} 应收敛向 μ 的特征函数. 已知 $f_{n_k} \to g$, 因此 μ 的特征函数就是 g. 因此 $\{\mu_n\}$ 的任一弱收敛的子序列的极限具有同一个特征函数. 由于特征函数唯一决定测度, 所以 $\{\mu_n\}$ 的任一弱收敛的子序列的极限相同. 由定理 8.5.13 知 $\mu_n \Longrightarrow \mu$ 且 μ 的特征函数为 g. □

6. 推论 若概率测度序列 $\{\mu_n\}$ 相应的特征函数序列 $f_n(u) \to g(u)$, 对一切 $u \in \mathbb{R}$ 成立, 且 $\{\mu_n\}$ 胎紧, 则存在概率测度 μ, 使得 $\mu_n \Longrightarrow \mu$ 且 μ 的特征函数为 g.

证明与推论 5 的后一部分类似.

习题

1. 试证 $\lim_{T\to\infty} \dfrac{1}{2T}\displaystyle\int_{-T}^{T} e^{-iux_0} f(u)\,du = \mu(\{x_0\})$.

2. 试证 $\lim_{T\to\infty} \dfrac{1}{2T}\displaystyle\int_{-T}^{T} |f(u)|^2\,du = \sum_{x\in\mathbb{R}} \mu(\{x\})^2$.

3. 若 X 的特征函数为 f, 概率分布为 μ, 则

$$f\text{是实函数} \Longleftrightarrow X\text{与}-X\text{同分布} \Longleftrightarrow \mu(B) = \mu(-B),\ \forall B \in \mathcal{B},$$

其中 $-B := \{-x : x \in B\}$.

4. 证明唯一性定理对于 $(\mathbb{R}, \mathcal{B})$ 上的有限符号测度也成立, 即若 μ, ν 是有限符号测度, 且

$$\int_{-\infty}^{\infty} e^{iux}\,\mu(dx) = \int_{-\infty}^{\infty} e^{iux}\,\nu(dx), \quad \forall u \in \mathbb{R},$$

则 $\mu = \nu$.

5. 设概率测度 $\mu_n(\{0\}) = \dfrac{1}{2} = \mu(\{n\})$, 对一切 $n \in \mathbb{N}$ 成立, 试讨论 $\{\mu_n\}$ 及其特征函数的收敛性.

6. 设 X_λ 是均值为 λ 的服从 Poisson 分布的随机变量, 试证: 当 $\lambda \to \infty$ 时, $\dfrac{X_\lambda - \lambda}{\sqrt{\lambda}}$ 按分布率收敛向标准正态分布.

§3. 中心极限定理

在自然科学, 工程技术和经济科学中, 经常遇到这样的量:

它由许多彼此不相干的随机因素的叠加而产生，而每一因素起的作用都不大. 我们要研究这类现象的规律性. 用概率论的语言来说，就是要研究一类随机变量的规律，它由许多独立随机变量的和组成，组成这个和的每一随机变量都非常地"小". 确切地说，要研究的是项数越来越多值越来越"小"的独立随机变量的和组成的序列的极限分布律，即所谓的中心极限问题.

1. 中心极限问题的一般提法：设对每一 $n \in \mathbb{N}$, $X_{nk}, k = 1,2,\cdots,k_n$, 是独立随机变量，令 $S_n = \sum_{k=1}^{k_n} X_{nk}$, $n \in \mathbb{N}$, 问

1) S_n 的一切可能的极限分布是哪些？

2) S_n 收敛向指定的分布的条件是什么？

如果对 X_{nk} 不加任何限制，则问题 1) 就没有任何实际意义. 因为对任何随机变量 X, 可以令 $X_{n1} = X$, $X_{nk} = 0, k = 2,\cdots,k_n$, 于是我们有

$$S_n \xrightarrow{\mathcal{D}} X.$$

根据问题的实际背景，人们永远假定 $\{X_{nk}\}$ 满足 "**一致可渐近忽略条件**"：

$$\text{(U)} \qquad \max_{1\leqslant k\leqslant k_n} \mathbf{P}(|X_{nk}| \geqslant \varepsilon) \to 0 (n \to \infty), \ \forall \varepsilon > 0.$$

已经证明在条件 (U) 之下，一切可能的极限分布律是所谓的无穷可分律. 对问题 2) 也找出一般的充分必要条件及向常见的分布律 – 退化律，正态律，Poisson 律、稳定律等分布律收敛的条件.

为了简单起见，同时也照顾到方法和问题的一般性，我们将在更强的条件下讨论中心极限问题：即若用 σ_{nk}^2 表示 X_{nk} 的方差，存在常数 $c > 0$ 使得

$$\text{(C)} \qquad \max_{1\leqslant k\leqslant k_n} \sigma_{nk}^2 \to 0 (n \to \infty); \quad \sup_{1\leqslant n<\infty} \sum_{k=1}^{k_n} \sigma_{nk}^2 \leqslant c < \infty.$$

条件 (C) 称为 **有界方差条件**.

如果每一 $\mathbf{E}X_{nk}=0$, 利用 Чебышев 不等式，对任意 $\varepsilon>0$,

$$\max_k \mathbf{P}(|X_{nk}|\geqslant\varepsilon)\leqslant\frac{1}{\varepsilon^2}\max_k\sigma_{nk}^2\to 0.$$

即条件 (C) 蕴含条件 (U).

为书写方便，我们假设 X_{nk} 已经中心化，即 $\mathbf{E}X_{nk}=0$, 用 P_{nk}, f_{nk} 分别表示其分布律和特征函数，用 $\prod\limits_k$, $\sum\limits_k$, $\max\limits_k$ 分别表示 $\prod\limits_k^{k_n}$, $\sum\limits_k^{k_n}$ 和 $\max\limits_{1\leqslant k\leqslant k_n}$; 极限过程如无特别声明，均指对 $n\to\infty$ 来取的.

由 X_{nk}, $k=1,2,\cdots,k_n$ 的独立性易知， $f_{S_n}(u)=\prod\limits_k f_{nk}(u)$, 而 S_n 依分布收敛的充分必要条件是其特征函数收敛. 而 $\prod\limits_k f_{nk}(u)$ 收敛等价于 $\exp\left\{\sum\limits_k\log f_{nk}(u)\right\}$(当表达式有意义时) 收敛. 这里，对复数 z, $\log z$ 表示 z 的对数的主值 (注意：一般地说，$\log a+\log b$ 未必等于 $\log(ab)$. 而且即便 $f(t)$ 连续且 $\log f(t)$ 有意义， $\log f(t)$ 也未必连续). 但是，对数和的收敛性不易讨论. 需要借助于以下引理进一步转化.

2. 引理 (比较引理) 若 $\{X_{nk}\}$ 满足条件 (C) 且对一切 n,k, $\mathbf{E}X_{nk}=0$, 则对每一固定的 u, 存在 n_u, 当 $n\geqslant n_u$ 时， $\log f_{nk}(u)$ 存在有限，且

$$\sum_k\left\{\log f_{nk}(u)-\big(f_{nk}(u)-1\big)\right\}\to 0,\quad \text{当 } n\to\infty \text{ 时}.$$

证明 由 $\mathbf{E}X_{nk}=0$ 及初等不等式 $|e^{ix}-ix-1|\leqslant\dfrac{x^2}{2}$, 对一切实数 x 成立可知

$$|f_{nk}(u)-1|=\left|\int_{-\infty}^{\infty}\left[e^{iux}-iux-1\right]P_{nk}(dx)\right|$$

$$\leqslant \int_{-\infty}^{\infty} \frac{(ux)^2}{2} P_{nk}(dx) \leqslant \frac{1}{2}\sigma_{nk}^2 u^2.$$

从条件 (C) 可得

$$\max_k |f_{nk}(u) - 1| \leqslant \frac{u^2}{2} \max_k \sigma_{nk}^2 \to 0$$

及

$$\text{(1)} \qquad \sum_k |f_{nk}(u) - 1| \leqslant \frac{u^2}{2} \sum_k \sigma_{nk}^2 \leqslant \frac{c}{2} u^2.$$

于是对固定的 u, 存在 n_u, 当 $n \geqslant n_u$ 时,

$$\max_k |f_{nk}(u) - 1| \leqslant \frac{1}{2}.$$

容易证明当复数 z 的模不大于 1/2 时, $\log(1+z)$ 有意义且

$$|\log(1+z) - z| \leqslant |z|^2.$$

于是有

$$\begin{aligned}
&\left|\sum_k \{\log f_{nk}(u) - (f_{nk}(u) - 1)\}\right| \\
= &\left|\sum_k \Big\{\log\big(1 + (f_{nk}(u) - 1)\big) - \big(f_{nk}(u) - 1\big)\Big\}\right| \\
\leqslant &\sum_k \left| f_{nk}(u) - 1 \right|^2 \\
\leqslant &\max_k |f_{nk}(u) - 1| \sum_k |f_{nk}(u) - 1| \to 0. \qquad \square
\end{aligned}$$

注 利用初等不等式

$$\left|\prod_{k=1}^n z_k - \prod_{k=1}^n w_k\right| \leqslant \sum_{k=1}^n |z_k - w_k|,$$

及 (1) 式可得

$$\text{(1)} \qquad \left|\prod_k f_{nk}(u) - 1\right| \leqslant \sum_k |f_{nk}(u) - 1| \leqslant \tfrac{c}{2} u^2.$$

于是只要 $\prod_k f_{nk}(u)$ 对一切 $u \in \mathbb{R}$ 收敛, 其极限函数一定在 $u = 0$ 连续, 因而是特征函数.

比较引理把讨论 $\prod_k f_{nk}(u)$ 的收敛性及求它的极限的问题转化为讨论 $\exp\{\sum_k (f_{nk}(u)-1)\}$ 的收敛性及求其极限的问题.

令

$$\psi_n(u)=\sum_k (f_{nk}(u)-1).$$

由于 $\mathbf{E}X_{nk}=\int x\,P_{nk}(dx)=0$. 所以

$$\begin{aligned}\psi_n(u)&=\sum_k \int (e^{iux}-1-iux)\frac{1}{x^2}x^2\,P_{nk}(dx)\\&=\int \frac{e^{iux}-1-iux}{x^2}x^2\sum_k P_{nk}(dx).\end{aligned}$$

令

$$\mu_n(B)=\int_B x^2\sum_k P_{nk}(dx),\quad B\in\mathcal{B}.$$

则 μ_n 是 $(\mathbb{R},\mathcal{B})$ 上的 L-S 测度, 且一致有界.

$$\mu_n(\mathbb{R})=\int_{\mathbb{R}} x^2\sum_k P_{nk}(dx)=\sum_k\int x^2\,P_{nk}(dx)=\sum_k \sigma_{nk}^2\leqslant c.$$

此外, μ_n 是关于 $\sum_k P_{nk}$ 的不定积分, 其 Radon 导数 $\dfrac{d\mu_n}{d\sum_k P_{nk}}=x^2$.

同时

$$\frac{e^{iux}-1-iux}{x^2}\to 0,\quad 当|x|\to\infty时,$$
$$\frac{e^{iux}-1-iux}{x^2}\to -\frac{u^2}{2},\quad 当|x|\to 0时.$$

因而 $\dfrac{e^{iux}-1-iux}{x^2}$ 是一有界连续函数 (对每一固定的 u), 因此关于 μ_n 可积. 由推论 7.2.7, 可以将 $\psi_n(u)$ 表示成下述形式:

(**) $$\psi_n(u)=\int\frac{(e^{iux}-1-iux)}{x^2}\,\mu_n(dx).$$

以下证明当 $e^{\psi_n(u)}$ 的极限存在时，极限应具有 $e^{\psi(u)}$ 的形式，其中：

$$(*)\qquad \psi(u)=\int\frac{(e^{iux}-1-iux)}{x^2}\,\mu(dx),$$

这里 μ 是有限 L-S 测度， $\mu_n\xrightarrow{v}\mu$ 且 $\mu(\mathbb{R})\leqslant c$. 从而满足条件 (C) 期望为 0 的独立和的极限分布的特征函数具有形式 $e^{\psi(u)}$. 为此先证明以下几个引理.

3. 引理 (*) 中的 ψ 与 μ 相互唯一决定.

证明 只需证 μ 由 ψ 唯一决定. 由控制收敛定理的推论 5.4.10 知

$$\psi'(u)=i\int\frac{e^{iux}-1}{x}\,\mu(dx)$$

$$\psi''(u)=-\int e^{iux}\,\mu(dx)$$

因此 $-\psi''(u)$ 是有限 L-S 测度 μ 的特征函数 (因为有限 L-S 测度 μ 除以常数 $\mu(\mathbb{R})$ 就是概率测度，其特征函数可以完全类似定义且具有完全类似的性质). 由定理 2.2 的类比， μ 由 $-\psi''$ 完全决定，从而被 ψ 所决定. □

4. 引理 沿用引理 2 的条件和记号. 对一切 $u\in\mathbb{R}$, $e^{\psi_n(u)}\to$ 某一$g(u)$ 的充分必要条件是存在有限 L-S 测度 μ 使得 $\mu_n\xrightarrow{v}\mu$. 此时 $g(u)=e^{\psi(u)}$, 其中 $\psi(u)$ 由 (*) 式决定.

证明 因为对给定的 u, $\dfrac{e^{iux}-1-iux}{x^2}$ 在 $\mathbb{R}$ 上有界连续且当 $x\to 0$ 时趋于 $-\frac{u^2}{2}$, 当 $x\to\infty$ 时趋于零，故由定理 8.5.4 知, $\mu_n\xrightarrow{v}\mu$ 蕴含 $\psi_n(u)\to\psi(u)$. 因而 $e^{\psi_n(u)}\to e^{\psi(u)}$. 条件的充分性获证.

以下往证条件的必要性. 设 $e^{\psi_n(u)}\to g(u)$, 则由 μ_n 一致有界及 Helly 选择定理的推论 (8.5.3) 知存在子序列 $\mu_{n_k}\xrightarrow{v}$ 某一μ, 于

是由充分性的证明知 $\psi_{n_k}(u) \to \psi(u)$. 因而

$$g(u) = \exp\left\{\int \frac{e^{iux}-1-iux}{x^2}\mu(dx)\right\}.$$

若 μ_n 的另一淡收敛的子序列 $\mu_{n'_k} \xrightarrow{v} \nu$, 则我们又有

$$g(u) = \exp\left\{\int \frac{e^{iux}-1-iux}{x^2}\nu(dx)\right\}.$$

于是

$$(1) \quad \int \frac{e^{iux}-1-iux}{x^2}\mu(dx) = i2k(u)\pi + \int \frac{e^{iux}-1-iux}{x^2}\nu(dx),$$

其中 $k(u)$ 为取整数值的函数.

由 (1) 知当 $u=0$ 时, $k(u)=0$. 注意到 (1) 式左端函数及右端第二个函数的连续性及 $k(u)$ 为整数值函数知 $k(u)\equiv 0$. 再由引理 3 即知 $\mu=\nu$.

由于 $\{\mu_n\}$ 的任意子序列都有淡收敛的子序列, 且淡收敛向同一 μ, 于是由推论 8.5.3 即得 $\mu_n \xrightarrow{v} \mu$. □

5. 引理 每一形如 $e^{\psi(u)}$ 的函数, 其中 $\psi(u)$ 如 (*) 所述, 都是一数学期望为零, 方差 $\sigma^2=\mu(\mathbb{R})$ 的随机变量的特征函数, 并且是一满足条件 (C) 的和的极限分布律的特征函数.

证明 首先, 由于

$$\begin{aligned}\psi(u) &= \int \frac{e^{iux}-1-iux}{x^2}\mu(dx)\\ &= \lim_{N\to\infty}\int_{[-N,N)} \frac{e^{iux}-1-iux}{x^2}\mu(dx),\end{aligned}$$

其被积函数对于给定的 u 是 $\mathbb{R}$ 上的有界连续函数, 因而可以通过将 $[-N,N)$ 分成小区间, 用阶梯函数的积分来逼近: 设分法 T 为

$$-N = x_{n0} < x_{n1} < \cdots < x_{nk_n} = N, \quad (0\text{不作分点}).$$

令

$$g_T(x) = \sum_{k=0}^{k_n-1} \frac{e^{iux_{nk}} - 1 - iux_{nk}}{x_{nk}^2} I_{[x_{nk}, x_{n(k+1)})}(x).$$

于是若令 $l(T)$ 表示分法 T 的最大子区间的长度，则

$$\int_{-N}^{N} \frac{e^{iux} - 1 - iux}{x^2} d\mu = \lim_{l(T)\to 0} \int_{[-N,N)} g_T(x)\, d\mu$$

$$= \lim_{l(T)\to 0} \sum_{k=0}^{k_n-1} \left[iua_{nk} + \lambda_{nk}(e^{iub_{nk}} - 1)\right],$$

其中 $\lambda_{nk} = \frac{1}{x_{nk}^2}\mu\big([x_{nk}, x_{n(k+1)})\big)$, $a_{nk} = -\lambda_{nk}x_{nk}$, $b_{nk} = x_{nk}$. 由此可见

$$\exp\left[\int_{-N}^{N} \frac{e^{iux} - 1 - iux}{x^2}\, \mu(dx)\right]$$

是 Poisson 型特征函数 (若 X 服从参数为 λ 的 Poisson 分布，则称 $a+bX$ 为 Poisson 型随机变量，其特征函数就是 $e^{iau+\lambda(e^{ibu}-1)}$) 乘积的极限，且易见它在零点连续，因而是特征函数. 因为 $e^{\psi(u)}$ 也是特征函数的极限且连续，所以也是一随机变量的特征函数. □

6. 推论 在条件 (C) 之下，若对一切 $u \in \mathbb{R}$,

$$\lim_{n\to\infty} \prod_k f_{nk}(u) = g(u) \tag{1}$$

存在，则 $g(u)$ 必为某一随机变量的特征函数. 且 $g(u) = e^{\psi(u)}$, 其中 $\psi(u)$ 具有 (*) 的形式.

证明 由 (1) 中极限存在，比较引理及 ψ_n 的定义知对任何 $u \in \mathbb{R}$,

$$\lim_{n\to\infty} \prod_k f_{nk}(u) = \lim_{n\to\infty} \exp\{\log \prod_k f_{nk}(u)\}$$

$$= \lim_{n\to\infty} \exp\{\sum_k \log f_{nk}(u)\}$$

$$= \lim_{n\to\infty} \exp\{\sum_k \log f_{nk}(u) - \psi_n(u)\}\exp\{\psi_n(u)\}$$
$$= \lim_{n\to\infty} e^{\psi_n(u)}.$$

再由引理 4 知 $g(u) = e^{\psi(u)}$, 其中 $\psi(u)$ 具有 (*) 的形式. 最后由引理 5 知 $g(u)$ 是随机变量的特征函数.

7. 引理 $e^{\psi_n(u)} \to e^{\psi(u)}$. 且 $\sum\limits_k \sigma_{nk}^2 \to \sigma^2$ 的充分必要条件是 $\mu_n \Longrightarrow \mu$, 此处 σ^2 是极限分布的方差.

证明 由 $\mu_n(\mathbb{R}) = \sum\limits_k \sigma_{nk}^2$, $\mu(\mathbb{R}) = i^{-2}\frac{d^2}{du^2}e^{\psi(u)}\Big|_{u=0} = \sigma^2$, 弱收敛的定义及引理 4 即得. □

现在我们给出有界方差条件下中心极限定理的完整叙述:

8. 定理 设对一切 $n \in \mathbb{N}$, $X_{nk}, k = 1, 2, \cdots, k_n$, 是独立随机变量, $\mathbf{E}X_{nk} = 0$, $DX_{nk} = \sigma_{nk}^2$, 且它们满足条件 (C). 则

1) 如果 $\prod\limits_k f_{nk}(u) \to g(u)$, 对一切 $u \in \mathbb{R}$, 则 $g(u)$ 必是特征函数且形如 $e^{\psi(u)}$, 其中 $\psi(u)$ 由 (*) 定义. 满足定理条件的随机变量和的一切可能的极限分布律的特征函数与一切形如 $e^{\psi(u)}$ 的函数类重合.

2) $\sum\limits_k X_{nk} \xrightarrow{\mathfrak{D}} X$, 且 $f_X(u) = e^{\psi(u)}$ 的充分必要条件是 $\mu_n \xrightarrow{v} \mu$, 其中

$$\mu_n(B) = \sum_k \int_B x^2\, P_{X_{nk}}(dx), \tag{1}$$

而 μ 与 $\psi(u)$ 由 (*) 式相互唯一决定.

3) $\sum\limits_k X_{nk} \xrightarrow{\mathfrak{D}} X$, $f_X(u) = e^{\psi(u)}$ 且 $\sum\limits_k \sigma_{nk}^2 \to DX$ 的充要条件是 (1) 中的 $\mu_n \Longrightarrow \mu$, 其中的 μ 与 $\psi(u)$ 由 (*) 式相互唯一决定.

对于非中心化的情形我们不加证明地给出下述定理. 有兴趣的读者可参考 [YWL] 第八章 §2 定理 2 及其引理.

9. 定理 设对一切 $n \in \mathbb{N}$, $X_{nk}, k = 1, 2, \cdots, k_n$. 是独立随机变量, $\mathbf{E}X_{nk} = a_{nk}$; $DX_{nk} = \sigma_{nk}^2$, 且它们满足条件 (C). 则

1) 如果 $\prod\limits_k f_{nk}(u) \to g(u)$, 对一切 $u \in \mathbb{R}$, 则 $g(u)$ 必是特征函数且形如 $e^{\widetilde{\psi}(u)}$, 其中

$$\widetilde{\psi}(u) = iau + \int \frac{e^{iux} - 1 - iux}{x^2} \widetilde{\mu}(dx). \tag{1}$$

满足定理条件的随机变量和的一切可能的极限分布律的特征函数与一切形如 $e^{\widetilde{\psi}(u)}$ 的函数类重合.

2) $\sum\limits_k X_{nk} \xrightarrow{\mathcal{D}} X$, 且 $f_X(u) = e^{\widetilde{\psi}(u)}$ ($\widetilde{\psi}(u)$ 形如 (1)) 的充分必要条件是 $\widetilde{\mu}_n \xrightarrow{v} \widetilde{\mu}$ 且 $\sum\limits_k a_{nk} \to a$, 其中

$$\widetilde{\mu}_n(B) = \sum_k \int_B x^2 P_{X_{nk} - a_{nk}}(dx), \tag{2}$$

而 $\widetilde{\psi}(u)$ 与 $(a, \widetilde{\mu})$ 由 (1) 式相互唯一决定.

3) $\sum\limits_k X_{nk} \xrightarrow{\mathcal{D}} X$, $f_X(u) = e^{\widetilde{\psi}(u)}$ 且 $\sum\limits_k \sigma_{nk}^2 \to DX$ 的充要条件是 (2) 中的 $\widetilde{\mu}_n \Longrightarrow \widetilde{\mu}$ 且 $\sum\limits_k a_{nk} \to a$. 其中的 $\widetilde{\psi}(u)$ 与 $(a, \widetilde{\mu})$ 由 (1) 相互唯一决定.

现在利用所得到的结果推出收敛于给定的分布律的条件.

10. 定理 若 $X_{nk}, k = 1, 2, \cdots, k_n$ 独立, ，$\mathbf{E}X_{nk} = 0$, $\sigma_{nk}^2 = DX_{nk}$ 有限, $k = 1, 2, \cdots, k_n$; $n \in \mathbb{N}$ 且 $\sum\limits_k \sigma_{nk}^2 = 1$(或 $\lim\limits_{n\to\infty} \sum\limits_k \sigma_{nk}^2 \to 1$), 则 $\sum\limits_k X_{nk} \xrightarrow{\mathcal{D}} N$ (此处及以下用 N 表示标准正态随机变量).

且 $\max\limits_k \sigma_{nk}^2 \to 0$ 的充分必要条件是

$$(1) \qquad g_n(\varepsilon) = \sum_k \int_{|x| \geqslant \varepsilon} x^2 P_{X_{nk}}(dx) \to 0.$$

证明 正态律 $N(0,1)$ 的特征函数是 $e^{-\frac{u^2}{2}}$, 即 $\psi(u) = -\frac{u^2}{2}$. 故由 (*) 知与 $\psi(u)$ 相应的 μ 为

$$\mu(\{0\}) = 1, \quad \mu(\mathbb{R} - \{0\}) = 0.$$

由于已知 $\mu_n(\mathbb{R}) = \sum\limits_k \sigma_{nk}^2 \to 1 = \mu(\mathbb{R})$. 故 $\sum\limits_k X_{nk} \xrightarrow{\mathfrak{D}} N$ 且 $\max\limits_k \sigma_{nk}^2 \to 0$ 的充分必要条件是 $\mu_n \xrightarrow{w} \mu$. 即对一切不以 0 为端点的开区间 I, 有 $\mu_n(I) \to \mu(I)$. 在此情况下它等价于

$$\sum_k \int_{|x|<\varepsilon} x^2 P_{X_{nk}}(dx) \to 1, \quad \forall \varepsilon > 0.$$

$$\sum_k \int_{\varepsilon_1 < |x| < \varepsilon_2} x^2 P_{X_{nk}}(dx) \to 0, \quad \forall\, 0 < \varepsilon_1 < \varepsilon_2.$$

由 $\sum\limits_k \sigma_{nk}^2 \to 1$, 这等价于 (1). 而 (1) 成立时,

$$\max_k \sigma_{nk}^2 = \max_k \int x^2 P_{X_{nk}}(dx) \leqslant \varepsilon^2 + \max_k \int_{|x| \geqslant \varepsilon} x^2 P_{X_{nk}}(dx)$$

对任意 $\varepsilon > 0$ 成立, 因而条件 (1) 蕴含着 $\max\limits_k \sigma_{nk}^2 \to 0$. □

特别, 设 $X_1, X_2, \cdots$ 是一独立随机变量序列, $\mathbf{E}X_n, DX_n$ 存在有限, $n = 1, 2, \cdots$, 对一切 n 取

$$X_{nk} = \frac{X_k - \mathbf{E}X_k}{s_n}, \ k = 1, 2, \cdots, n, \ s_n^2 = \sum_{k=1}^n DX_k.$$

显然它们满足定理 10 的条件, 用 F_k 表示 X_k 的分布函数, 则 X_{nk} 的分布函数 $F_{nk}(x) = F_k(s_n x + \mathbf{E}X_k)$, 于是

$$g_n(\varepsilon) = \sum_{k=1}^n \int_{|x| \geqslant \varepsilon} x^2 \, dF_{nk}$$

$$= \sum_{k=1}^{n} \int_{\left|\frac{x-EX_k}{s_n}\right| \geqslant \varepsilon} \left| \frac{x - EX_k}{s_n} \right|^2 dF_k$$

$$= \frac{1}{s_n^2} \sum_{k=1}^{n} \int_{|x-EX_k| \geqslant \varepsilon s_n} |x - \mathbf{E}X_k|^2 \, dF_k$$

$$\leqslant \frac{1}{\varepsilon^\delta s_n^{2+\delta}} \sum_{k=1}^{n} \int |x - \mathbf{E}X_k|^{2+\delta} \, dF_k$$

$$= \frac{1}{\varepsilon^\delta s_n^{2+\delta}} \sum_{k=1}^{n} \mathbf{E}|x - \mathbf{E}X_k|^{2+\delta},$$

其中 δ 是任意正数. 由此可得下述二推论.

11. 定理 (**Lindeberg-Feller 定理**) 设 $\{X_n\}$ 是独立随机变量序列, $\mathbf{E}X_n, DX_n$ 存在有限, 则

$$\frac{\sum_{k=1}^{n}(X_k - EX_k)}{s_n} \xrightarrow{\mathfrak{D}} N \text{及} \max_k \frac{\sigma_k}{s_n} \to 0$$

的充分必要条件是对一切 $\varepsilon > 0$, 当 $n \to \infty$ 时,

$$g_n(\varepsilon) = \frac{1}{s_n^2} \sum_{k=1}^{n} \int_{|x-EX_k| \geqslant \varepsilon s_n} |x - \mathbf{E}X_k|^2 \, dP_{X_k} \to 0.$$

上述条件称为 Lindeberg 条件.

12. 定理 (**Ляпунов 定理**) 设 $\{X_n\}$ 是独立随机变量序列, $\mathbf{E}X_n, DX_n$ 存在有限, 且对某一 $\delta > 0$, 当 $n \to \infty$ 时,

$$\frac{1}{s_n^{2+\delta}} \sum_{k=1}^{n} \mathbf{E}|X_k - \mathbf{E}X_k|^{2+\delta} \to 0,$$

则

$$\frac{\sum_{k=1}^{n}(X_k - EX_k)}{s_n} \xrightarrow{\mathfrak{D}} N.$$

13. 推论 若 $X_{nk}, k = 1, 2, \cdots, k_n$ 独立, $\mathbf{E}X_{nk} = a_{nk}$, $DX_{nk} = \sigma_{nk}^2$ 有限, $k = 1, 2, \cdots, k_n$, $n \in \mathbb{N}$, 则 $\sum_k X_{nk} \xrightarrow{\mathfrak{D}} a(a$

为一实数) 的充分必要条件是 $\sum\limits_k a_{nk} \to a$ 且

$$\sum_k \int_{|x-a_{nk}|\geqslant\varepsilon} |x-a_{nk}|^2 \, d\mathbf{P}_{nk} \to 0, \ \forall\ \varepsilon > 0.$$

证明留作习题.

14. 定理 若 $X_{nk}, k=1,2,\cdots,k_n$ 独立, , $\mathbf{E}X_{nk}=a_{nk}$ 和 $DX_{nk}=\sigma_{nk}^2$ 有限, $k=1,2,\cdots,k_n$, $n\in\mathbb{N}$ 且 $\max\limits_k \sigma_{nk}^2 \to 0, \sum\limits_k \sigma_{nk}^2 \to \lambda$, 则 $\sum\limits_k X_{nk} \xrightarrow{\mathcal{D}} \mathcal{P}(\lambda)$($\mathcal{P}(\lambda)$ 表示参数为 λ 的 Poisson 随机变量) 的充分必要条件是 $\sum\limits_k \mathbf{E}X_{nk} \to \lambda$ 且对任一 $\varepsilon > 0$,

$$\sum_k \int_{|x-\lambda|\geqslant\varepsilon} x^2 \, dP_{X_{nk}-EX_{nk}} \to 0.$$

证明留作习题.

最后, 我们对满足条件 (U) 的中心极限定理的一般结果作一简单介绍.

15. 定义 一个分布律称为 **无穷可分律**, 如果它的特征函数 f 具有性质: 对每一 n, 存在一个特征函数 f_n, 使得 $f=f_n^n$. 这时也称特征函数 f 是 **无穷可分** 的.

16. 定理 任何有限个无穷可分的特征函数的乘积仍是无穷可分的. 无穷可分律的极限分布仍是无穷可分律.

17. 定理 每一无穷可分的特征函数均具有 $e^{\psi(u)}$ 的形式, 其中

$$\psi(u) = iua + \int \left(e^{iux} - 1 - \frac{iux}{1+x^2}\right) \frac{1+x^2}{x^2}\, \mu(dx), \tag{1}$$

其中 μ 为 $\mathbb{R}$ 上有限 L-S 测度, $a\in\mathbb{R}$. 且 ψ 与 (a,μ) 相互唯一决定 (记作 $\psi=(a,\mu)$).

18. 定理 (**中心极限定理**) 设 $\{X_{nk}\}$ 满足条件 (U), 则

1) $\sum\limits_k X_{nk}$ 的一切可能的极限分布是一切无穷可分律;

2) $\sum\limits_k X_{nk} \xrightarrow{\mathcal{D}} X$, $f_X(u) = e^{\psi(u)}$, 其中 $\psi = (a,\mu)$(为定理 17 所述), 的充分必要条件是

$$a_n \to a, \quad \mu_n \xrightarrow{v} \mu,$$

其中

$$a_n = \sum_k \left\{ a_{nk} + \int \frac{x}{1+x^2}\, dP_{X_{nk}-a_{nk}} \right\},$$

$$\mu_n(B) = \sum_k \int_B \frac{x^2}{1+x^2}\, dP_{X_{nk}-a_{nk}}, \quad B \in \mathcal{B}.$$

而

$$a_{nk} = \int_{|x|<\tau} x\, dP_{X_{nk}},$$

τ 是任意取定的一个正数.

习题

1. 证明推论 13, 14.

2. 若 X_n 在 $[-n,n]$ 上均匀分布, $n \in \mathbb{N}$, 试证 $\{X_n\}$ 满足 Lindeberg 条件.

3. 设 $X_1, X_2, \cdots$ 是独立随机变量序列, $S_n = \sum\limits_{k=1}^{n} X_k$, $s_n^2 = \sum\limits_{k=1}^{n} DX_k$, 判断在下述哪种情况下

$$\frac{S_n - \mathbf{E}S_n}{s_n} \xrightarrow{\mathcal{D}} N.$$

i) $\mathbf{P}(X_k = -2^k) = \mathbf{P}(X_k = 2^k) = \frac{1}{2}$, $k = 1, 2, \cdots$;

ii) $\mathbf{P}(X_k = -2^k) = \mathbf{P}(X_k = 2^k) = 2^{-(k+1)}$,

$\mathbf{P}(X_k = 0) = 1 - 2^{-k}$, $k = 1, 2, \cdots$;

iii) $\mathbf{P}(X_k = k) = \mathbf{P}(X_k = -k) = \frac{1}{2}k^{-\frac{1}{2}}$,

$\mathbf{P}(X_k = 0) = 1 - k^{-\frac{1}{2}}$, $k = 1, 2, \cdots$.

4. 设 X_k, $k = 1, 2, \cdots$ 为 iid, X_1 是参数为 1 的 Poisson 随机变量, 应用 Ляпунов 定理到这一序列, 证明

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

5. 证明形如 $e^{\psi(u)}$ (其中 $\psi(u)$ 为 (*) 所示) 的特征函数是无穷可分的.

6. 设特征函数 f 是无穷可分的, 试证对一切 u, $f(u) \neq 0$. (提示: 考虑 $g(u) = |f(u)|^2$, 证明它仍是无穷可分的, 且 $g(u)$ 恒不为零.)

参 考 文 献

[B1] Billingsley, P., Probability and measure, Wiley, 1979.

[B2] Billingsley, P., Convergence of Probability measures, John Wiley & sons, 1968.

[Br] Breiman, L., Probability, Addison-Wesley, 1968.

[Ch1] Chung, K.L., A course in probability theory, Academic Press, 1974 (中译本: 钟开莱著, 刘文, 吴让泉译, 概率论教程, 上海科学技术出版社, 1989).

[Ch2] Chung, K.L., Elementary probability theory with stochastic processes, Springer–Verlag, 1974 (中译本:钟开莱著,魏宗舒等译,初等概率论附随机过程, 人民教育出版社, 1979).

[Co] Cohn, D.L., Measure theory, Birhäuse Boston, 1980.

[Cr] Cramer, H., Mathematical methods of statistics,Princeton Univ. Press, 1946(中译本: H. 克拉美著, 魏宗舒等译, 统计学数学方法, 上海科学技术出版社, 1960).

[CT] Chow, Y.S. and Teicher, H., Probability Theory, Springer–Verlag, 1988.

[Di] Dieudonne, J., Foundations of modern analysis, enlarged and correted printing, Academic press, 1969 (中译本: J. 迪厄多内著, 郭瑞芝, 苏维宜译, 科学出版社, 1982).

[Do] Doob, J.L., Stochastic processes, Wiley, 1953.

[Du] Durret,R., Probability: theory and examples, Wadsworth &Broks /Cole, 1991.

[Fa1] Falconer, K.J., Fractal geometry–mathematical foundations and applications, 1990 (中译本: 肯尼思. 法尔科内著, 曾文曲, 刘世耀等译, 分形几何 – 数学基础及其应用, 东北工学院出版社, 1991).

[Fa2] K.Falconer, The geometry of Fractal sets, Cambridge university press, 1985.

[Fe] Feller, W., An introduction to probability theory and its applications, 1–2, Wiley, 1957–1971 (第一卷中译本: 威廉. 费勒著, 上册胡迪鹤, 林向清译, 下册刘文译, 概率论及其应用, 科学出版社, 1964,1979).

[Fr] Freedman, D., Markov chains, Springer–Verlag, 1983.

[GK] Гнеденко, Б.В. и Колмогоров, А.Н., Пределйные распределения, длясумм независимых случайных величин, Гостеиздат, 1949 (中译本: Б.В. 格涅坚科和 А.Н. 廓洛莫格若夫著, 王寿仁译, 相互独立随机变数之和的极限分布, 科学出版社, 1955).

[GS] Гихман, И.И., Скороход, А.В., Теория случайных процессов, 2, М., 1973 (英译本: Gihman, I.I. and Skorohod, A.V., Theory of stochastic processes, 2, Springer, 1979; 中译本: 随机过程论, 科学出版社, 1986).

[Ha] Halmos, P. R., Measure theory, Van Nostrand, 1950 (中译本: 王建华译, 测度论, 科学出版社, 1958).

[Hu] 胡迪鹤, 分析概率论, 科学出版社, 1984.

[HWY] He,S.W., Wang,J.G., Yan,J.A., Semimartingale theory and stochastic calculus, Science Press and CRC Press Inc., 1992.

[Ka1] Kac, M., Probability and related topics in physical sciences, Wiley, 1959.

[Ka2] Kac, M., Statistical independence in probability, analysis and number theory, Math. Assoc. Amer., 1963.

[Ko1] Колмогоров, А.Н., Основные понятия теории вероятностей, 2 изд.,М., 1974 (英译本: Kolmogorov, A.N., Fundations of the theory of probability, Chelsea, reprint, 1950; 中译本: А.Н. 柯尔莫格洛夫著, 丁寿田译, 概率论基本概念, 商务印书馆, 1952).

[Ko2] Колмогоров,А.Н., Теория вероятностей, в кн., Математика, ее содержание. методы и значения, М.,1956 (英译本: Kolmogorov, A.N., The theory of probability, in *Mathematics, its content, methods and mcaning*, Vol.4, Amer. Math. Soc., 1963).

[LLL] 陆传荣, 林正炎, 陆传赉, 概率论极限理论引论, 高等教育出版社, 1989.

[Lo] Loéve, M., Probability theory, Springer–Verlag, 1978 (中译本: M. 洛易甫著, 梁文骐译, 概率论, 上册, 科学出版社, 1965).

[LR] Laha, R.G. and Rohatgi, V.K., Probability theory, John Wiley & sons, 1979.

[Ne1] Neveu, J., Discrete-parameter martingales, North–Holland, 1975 (译自法文).

[Ne2] Neveu, J., Mathematical fundations of the calculus of probability, North–Holland, 1965 (译自法文).

[Pa1] Parthasarathy, K.R., Introduction to probability and measure, Macmillan, 1977.

[Pa2] Parthasarathy, K.R., Probability measures on metric spaces, Academic Press, 1967.

[Paz] Parzen, E., Stochastic processes, Holden–Day, 1962 (中译本：邓永录，杨振明译，伊曼纽尔·帕尔逊著，随机过程，高等教育出版社，1987).

[QL] 钱佩玲，柳藩，实变函数论，北京师范大学出版社，1991.

[Re] Rényi, A., Foundations of probability, Holden–Day, 1970.

[Ro] Rosenblatt, M., Random processes, Springer, 1974.

[Sh] Shiryayev, A.N., Probability, Springer–Verlag, 1984.

[SS] Steen, L.A., Seebach, J.A.Jr., Counter examples in Topology, Springer-Verlag, 1978.

[Su] 孙永生，泛函分析讲义，北京师范大学出版社，1986.

[We] Вентцель А.Д., Курс теории случайных процессов, Наука, 1976. (英译本： Wentzell, A.D., A course in the theory of stochastic processes, McGraw-Hill, 1981.)

[Wj] 汪嘉冈，现代概率论基础，复旦大学出版社，1988.

[Ws] 王寿仁，概率论基础和随机过程，科学出版社，1986.

[Wz] 王梓坤，随机过程论，科学出版社，1965.

[Xi] Хинчин, А.Я., Математические методы теории массового обслуживания, Изд, АХ СССР, 1955 (中译本： А.Я. 欣钦著，张里千，殷涌泉译，公用事业理论的数学方法，科学出版社，1958).

[Zs] 中山大学“测度与概率基础”编写组，测度与概率基础，广东科技出版社，1984.

名 词 索 引

B

C

---

xt 表示习题

K

L

M

N

P

Q

R

S